JOURNAL *of* EDUCATIONAL RESEARCH

Volume XII DECEMBER, 1925 Number 5

THE EFFECT OF OBJECTIVE STANDARDS UPON COMPOSITION TEACHERS' JUDGMENTS

EARL HUDELSON
University of Minnesota

The whole policy of insisting upon definite standards of pupil achievement implies an ability on the part of teachers to determine when those standards have been attained. Subjective standards of pupil proficiency in English composition are automatically disqualified. Finite minds have amply demonstrated their inability to agree satisfactorily upon standards expressed subjectively. Objective standards offer difficulty enough; but only as they are derived for the various aspects of achievement in English composition will standards in that field be meaningful.

Figure 1 reveals typical conditions under the subjective régime. It shows the average percentage scores and the average deviations on seventeen compositions by seventeen experienced English teachers in the same school who had up to that time been satisfied with their ability to judge composition merit subjectively. The length of the vertical black lines indicates the average deviations of the seventeen teachers' scores from the average score of all seventeen teachers. The average scores are shown by short solid cross lines. The distance between each pair of dotted lines represents the average annual pupil progress in composition. The half of the judges who agreed most closely upon the value of theme *1* varied on the average 9.8 percent or over four composition years. Variations on several of the other themes are almost as great, that on theme *4* alone being as small as twice the average annual pupil progress. These seventeen members of the English faculty of one school not only differed significantly,

as was found later from the standard of teachers in other schools, but they disagreed widely among themselves.[1] If these teachers had essayed to classify the author of theme 5 on the basis of his score on this one theme, his placement would have differed by eight years according to two of the teachers and more than four years according to half of them.

The chaotic conditions that prevail under subjective standards

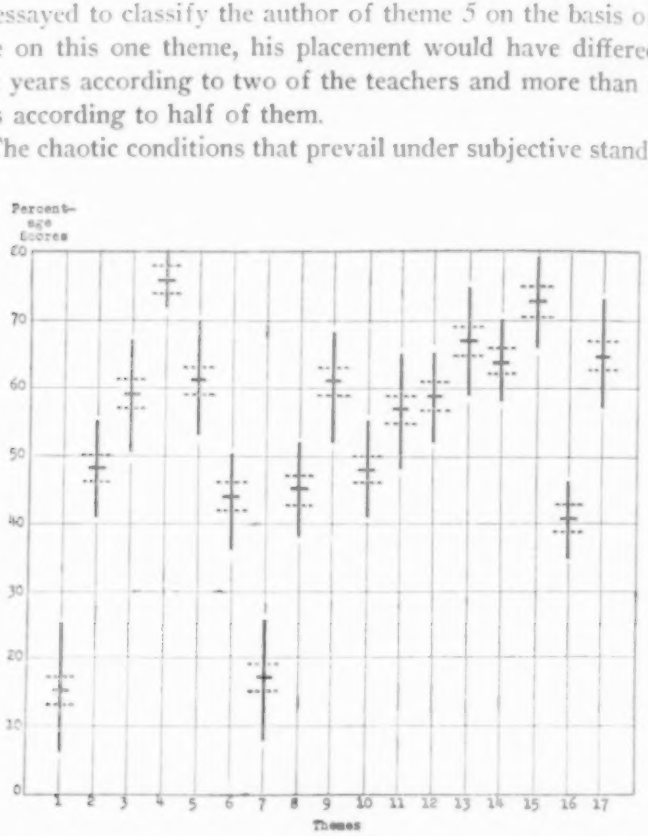

FIGURE 1. PERCENTAGE SCORES OF SEVENTEEN TEACHERS ON SEVENTEEN THEMES. VERTICAL BARS SHOW AVERAGE DEVIATION

are revealed again in Table I, which shows the percentage marks of ten teachers in the same school on eight compositions selected at random. Theme *713* received a percentage score of 40 from one teacher and 95 from another. Teacher *B* would have

[1] The range of scores on each of the seventeen themes was, respectively, 0-35, 35-60, 40-70, 70-82, 30-70, 30-60, 0-30, 30-75, 45-80, 30-60, 40-75, 48-80, 45-80, 50-83, 50-85, 27-50, and 45-80.

"passed" all eight of these pupils whereas teacher *A* would have "passed" only two of them, even though these two instructors were at that time teaching different sections of the same composition course in the same school. The pupil who wrote *713* was given by the six teachers who would have "passed" him an average score of 83 percent, while the four teachers who would have "failed" him gave him an average mark of 56 percent. No wonder pupils like to pick their instructors!

On the percentage scale the average annual pupil improvement is about 4 percent. On more than half of the themes represented in Table I, the average deviation of these ten teachers

TABLE I. PERCENTAGE SCORES OF EIGHT THEMES BY TEN TEACHERS IN SAME SCHOOL

Theme Number	Teachers										Average	Average Deviation
	A	B	C	D	E	F	G	H	I	J		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
706	55	88	71	80	60	90	86	85	90	70	77.5	±10.8
709	70	89	78	90	83	94	85	85	92	70	83.6	± 6.7
713	40	84	85	76	50	95	82	70	75	65	72.2	±12.7
726	80	94	75	92	80	88	91	90	95	90	87.5	± 5.5
740	75	87	79	90	70	70	81	79	85	75	79.1	± 5.3
761	40	92	72	86	75	92	80	80	60	75	75.2	±10.8
762	25	75	50	42	30	45	40	40	30	35	41.2	± 9.4
794	30	75	62	38	40	45	45	35	45	40	45.5	± 9.2

(*see* column *13*) was equivalent to more than two years of pupil improvement. On theme *713* it was equivalent to more than three years. On certain themes some of these teachers varied from the average judgment of all ten teachers as much as eight years. On theme *713* teacher *A* and teacher *F* were thirteen composition years apart in their scores in spite of the fact that they taught in the same building. This is not a case of a group of teachers differing from another group or from a national standard; it is a case of wide disagreement among themselves.

The significant fact about these results is not that those teachers were hopeless judges, for practice with objective standards soon revealed that they were not; but that their pupils were

not getting justice under subjective standards, and that it took practice in the use of an objective scale to discover the injustices.

What effect do objective standards have upon the ability of English teachers to agree about composition quality? Figures 2, 3, and 4 reveal the progress made by three teachers in the same school in agreeing upon the ranking of ten compositions.[2] Figure 2 shows how widely these three teachers disagreed when they

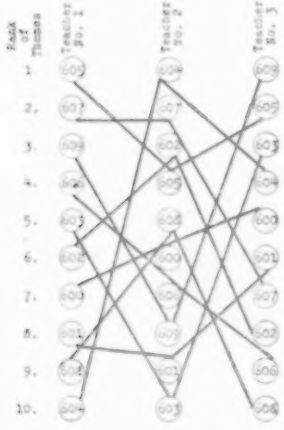

FIGURE 2. RANKINGS OF TEN THEMES BY THREE TEACHERS WITHOUT THE USE OF A SCALE

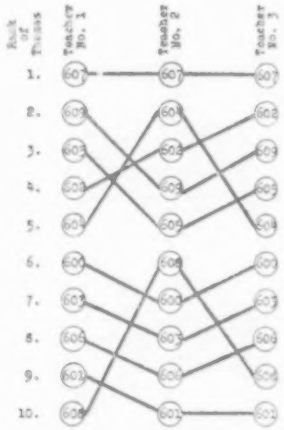

FIGURE 3. RANKINGS OF THE SAME TEN THEMES BY THE SAME THREE TEACHERS ONE MONTH LATER WITH THE USE OF A SCALE

ranked these ten themes by subjective standards. It will be noted that theme *604* was ranked first by one teacher, fourth by another, and last by the other. The standards were so chaotic that not even general tendencies are apparent; each teacher was higher than the other two about as often as she was lower. It is obvious in Figure 3 that the use of a scale one month later produced considerably more agreement and decidedly more uniformity. The second teacher was low, but she was consistently

[2] Hudelson, Earl. "English Composition: Its Aims, Methods, and Measurement." *Twenty-second Yearbook of the National Society for the Study of Education.* Bloomington, Illinois, Public School Publishing Company, 1923. Part I. pp. 32-34.

low. She could correct her error on practically all themes by correcting it on one. That is just what she did during the ensuing fourteen weeks of practice and conferences, with the result shown in Figure 4.

The writer has invariably found among English teachers that an erratic judgment of composition merit is harder to tame than a consistently misplaced judgment is to correct. In other words,

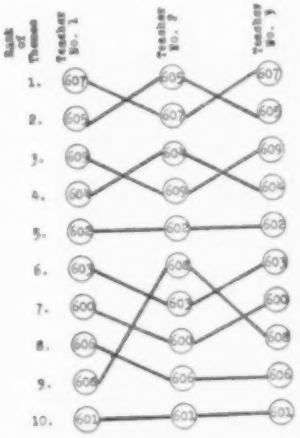

FIGURE 4. RANKINGS OF THE SAME TEN THEMES BY THE SAME THREE TEACHERS FOURTEEN WEEKS LATER WITH THE USE OF THE SAME SCALE

English teachers find it harder to reduce their variabilities of judgment than to correct their systematic errors in judgment. Figures 3 and 4 show the typical effect of objective standards upon misplaced judgment; data following will show that objective standards are capable of producing a no less salutary effect upon erratic judgments.

Table II shows the comparative percentage distribution of scores on six hundred themes scored by eight teachers without and with the use of an objective composition scale. The scale scores tend to conform more closely to the distribution which would be expected in heterogeneous class groups such as these

were. In every such analysis the writer has found that the use of a scale tends (1) to produce more nearly normal distributions of marks; (2) to decrease the frequency of superior marks; (3) to shift a considerable number of barely passing marks down into the failure level; and (4) to cause the class medians to rise at a more reasonable rate.

This last tendency is shown in Figure 5. Each of 408 pupils in grades IV to XII wrote upon 32 theme assignments during the course of one year. Each theme was scored twice by eight

TABLE II. COMPARISON OF THE PERCENTILE DISTRIBUTION OF COMPOSITION SCORES WITH AND WITHOUT USE OF SCALE (600 CASES)

Scores	Percentile Distributions							
	Grade VII		Grade VIII		Grade IX		Grade X	
	Without Scale	With Scale	Without Scale	With Scale	Without Scale	With Scale	Without Scale	With Scale
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
90–100	18.0	4.2	7.3	8.9	11.6	2.0	22.0	7.0
80– 89	36.6	20.6	18.0	22.0	20.0	26.0	18.8	21.0
70– 79	24.4	22.0	24.0	29.0	20.0	40.0	19.2	42.8
60– 69	6.3	13.0	13.5	21.7	15.7	16.8	11.1	15.2
50– 59	2.1	16.0	15.7	10.0	14.0	6.0	3.9	6.1
40– 49	5.3	11.0	15.0	2.1	6.2	4.2	5.0	3.9
30– 39	4.2	10.7	2.1	2.1	5.2	3.0	0.0	3.0
20– 29	3.1	3.1	1.1	4.2	7.3	2.0	0.0	1.0
10– 19	0.0	0.0	3.3	0.0	0.0	0.0	0.0	0.0
0– 9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

teachers working independently, first on a percentage basis and later with the use of an objective scale. Figure 5 shows the median judgment of the eight teachers on each scoring. It seems reasonable to expect improvement in composition upon the part of some of the pupils during each school year. Except for a slight negative acceleration while pupils are passing from a lower school into the next higher and becoming adjusted to the new, it would seem reasonable to anticipate comparatively uniform progress. If the percentage curve in Figure 5 is reliable, then

either children in grades VIII and XI lost some composition ability which they had already acquired or there is consistently poorer teaching in certain years. It is hardly reasonable to expect pupils who remain in school to backslide in their ability to write, and there is no cause for believing that all poor teaching gravitates

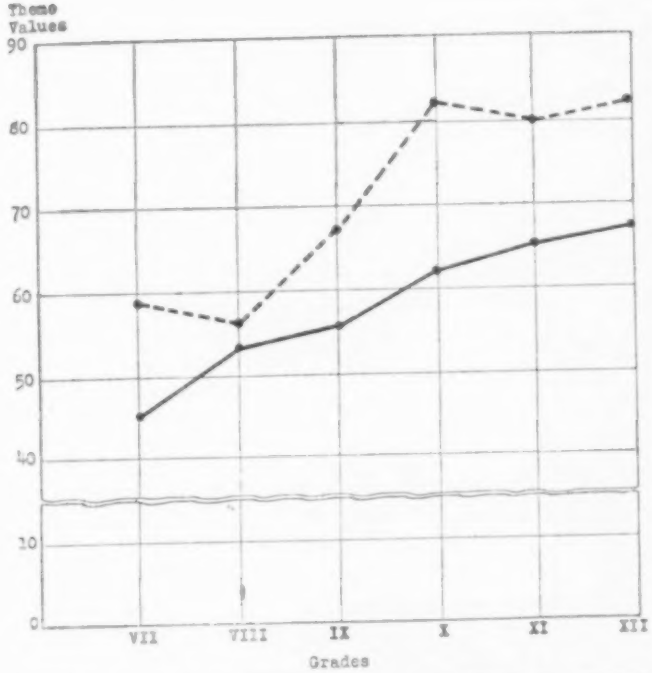

FIGURE 5. MEDIAN JUDGMENT OF EIGHT TEACHERS ON THEMES WRITTEN BY 408 PUPILS. SOLID LINE REPRESENTS SCALE SCORES; BROKEN LINE REPRESENTS PERCENTAGE SCORES

toward particular school grades. Progress in terms of scale scores in Figure 5 more nearly approximates expectancy.

Upon refinement of judgment Figure 6 shows the effect caused by practice in using a composition scale. It represents the experiences of 157 English teachers whom the writer trained to use an objective composition scale. As many pupil themes of unknown value were provided as there were teachers in the

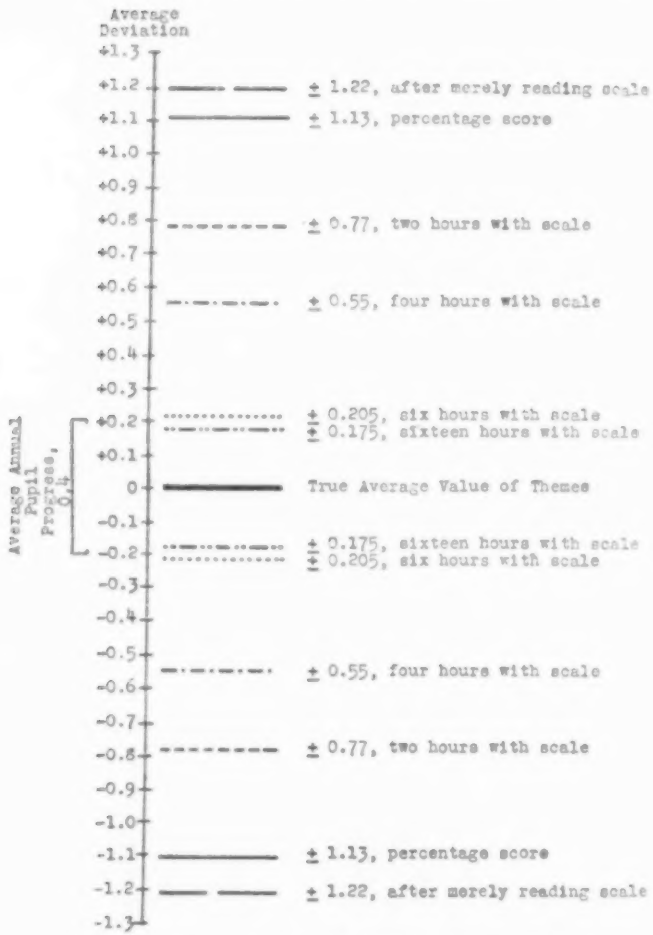

FIGURE 6. THE EFFECT UPON RELIABILITY OF PRACTICE IN USING THE HUDELSON TYPICAL COMPOSITION SCALE. 157 JUDGES AND 157 THEMES

practice groups. Each group, ranging in size from 28 to 57, scored the papers in round-robin fashion. Before the scale was introduced each teacher scored the themes in terms of percents. Next the teachers read the composition scale silently and then scored the themes by it. At the third meeting the scale was read

aloud by the instructor and discussed by the group, after which the teachers, always working independently, scored the themes again. After four, six, and sixteen hours of reading, discussion, and home practice the teachers rescored the themes. After each scoring the instructor collected the score sheets, and the teachers were asked not to communicate any scores which they remembered having given. They were not told the true values of the themes until after the scoring which followed sixteen hours of practice. They were given no clue as to whether their judgments were improving or not. It is reasonable to expect that improvement would have been more rapid if a clue had been given; but this "blind" procedure was followed in order to counteract the effect of their remembering how they had scored a theme previously.

The heavy black line in the middle of Figure 6 represents the true average value of the themes which were scored. The brace at the left spans the average annual amount of pupil-progress in composition. The distance from the middle line to the second horizontal line from either end represents the average of these 157 teachers' average deviations from the true scores when they scored the themes on a percentage basis. Half of these teachers, it will be seen, scored these themes over five years above or below their true values. The deviations were even greater immediately after the teachers had merely read the scale. Two hours of practice with the scale, however, brought them over a year closer to the true values. Four more hours of practice brought them a half-year closer, and two additional hours brought them another composition year nearer the actual values. After the sixth hour of practice their rate of improvement in reliability decreased rapidly; but with only sixteen hours of "blind" practice these teachers, who had had no previous experience with objective composition scales, were, on the average, scoring themes within a composition year of their true values. They were from two to five times as reliable in their judgments as are English teachers the country over who use only subjective standards. Rate of improvement would have de-

creased still more with each additional hour of practice; but there is ample evidence here to justify the use of objective scales in scoring compositions and to show that ability to use such scales with reasonable precision is not a forbiddingly slow process.

Composition scales, then, tend to refine judgment as well as to stabilize it. They perform the paradoxical function of rendering judgment both more constant and more amenable to improvement. Moreover, they tend to give teachers a more modest opinion of their ability to judge composition quality. Upon first using objective composition scales many teachers hold the same exalted estimate of their powers of discrimination that they held about their ability to measure composition merit subjectively. They insist upon giving themes such absurdly nice scale values as 5.68 or even 5.681. It is at first a disconcerting but eventually a wholesome revelation to them to discover that only three out of four trained judges can detect the average amount of pupil improvement in writing over a period of two-and-a-half years. They come to appreciate that progress in such a complex accomplishment as composition is very slow and that they need not chafe if their results over a single year seem to be meager. They do not seize upon this fact as an alibi for poor teaching, however, for they realize that they will have to teach at top speed to produce even this slight progress. What is even more important, teachers who learn to use objective composition scales realize that they have not been measuring their words when they declared that one pupil theme is ten times, or even twice, as good as another pupil theme. As soon as they come to comprehend the zero point of composition merit and to see concrete samples of multiple units of value they appreciate the fact that their former estimates of differences may have been exaggerated. They discover that Thorndike was right when he contended[3] that:

> The paragraph writing of pupils in our high schools and that of the world's hundred best English writers undoubtedly overlap considerably in merit. . . . In the judgment of high-school teachers of English, the worst tenth of paragraph writing of high-school pupils is nearly

[3] Thorndike, E. L. "Notes on the Significance and Use of the Hillegas Scale," *English Journal*, 2:551-61, November, 1913.

half-way from zero toward the best the world knows. What we rightly consider a mediocre composition still represents nearly three-fourths of the progress possible. . . . The world pays enormous sums in money and fame for a difference so small that one person out of four cannot see it.

The values which accrue from the use of objective composition standards are, on the whole, greatest for teachers whose powers of discrimination are already most refined. The writer has generally found that the composition judgment of those English teachers whose subjective standards are most reliable improves most and fastest in practice with objective standards. This was clearly demonstrated, for instance, by the English teachers of Central High School, St. Paul. For two years these teachers carried on a campaign to reduce variation in their subjective estimates of pupil-composition merit. At the end of that time their ability to agree was almost uncanny. Their first attempt to score by an objective standard was no better than that of other teachers; but after five-and-a-half hours of practice with a scale they were in practical agreement.

There is evidence that ability to use general merit composition scales reliably is a fair measure of the same factors of intelligence that our mental tests now measure. The writer ascertained the intelligence quotients of 91 English teachers. After they had practiced four hours in using a composition scale their average deviations correlated with their intelligence quotients — .69. After sixteen hours of "blind" practice with the scale the coefficient of correlation was — .735.

It would seem, then, that practice in the use of objective standards of pupil achievement in English composition renders the teachers' judgments more accurate and constant at the same time that it makes them more receptive and sensitive to refinement. It enables teachers to classify pupils more confidently and to detect composition improvement more reliably, and appears to give English teachers moral support in assigning marks justly. The use of objective measures tends to have a rationalizing influence upon English teachers, often revealing to them a presumptuously high opinion of their powers of discrimination

and an exaggerated sense of differences in composition merit, and curbs erratic judgments as well as corrects consistent errors of judgment. Reliability in the use of objective composition scales furnishes some indication of intelligence. Teachers whose accuracy on subjective standards of composition quality is highest generally learn to use objective standards most rapidly and reliably.

The quality of the available composition tests and the possibilities which these objective standards offer for improving English teachers' ability to measure pupils' composition achievement are out of proportion to the use that is being made of them. We wear braces of various sorts to support weak bodily structures; we plague the grocer for the use of his scales to scare our indulgent wills or pampered appetites into subjugation; we rely upon all sorts of devices for reinforcing our treacherous memories; but we hesitate to employ methods of supporting, reinforcing, and refining our perfectly normal fallibilities and limitations of human judgment. The use of objective devices in education is a contribution to meaningful standardization and is a step in professional self-improvement which any group of teachers can take and which sooner or later every professionally minded group will take. There are already enough open-minded teachers receptive to scientific means of social control to assure the success of educational measurements, whether the rest embrace them or not. All of the early remonstrances against the automobile availed nothing; automobiles came anyway because they afforded an improved means of social control. So will objective educational measurements come anyway; and, like the automobile, they will divide humanity into the quick and the dead.

A PLAN OF CLASSIFYING PUPILS

E. D. PRICE
Superintendent of Schools, Enid, Oklahoma

In 1918 the authorities of the Enid[1] public schools became interested in the idea of the classification of pupils according to mental ability. The superintendent set himself the task of planning for the necessary instruction of the teaching force and the needed reorganization. Two years were occupied in preparation; and during the time every teacher in the school system became proficient in the use of individual mental tests. All pupils from grades I to IX were tested by the Stanford Revision of the Binet-Simon Test; those from grades IX to XII, by group tests. The entire plan of preparation and organization was published in a bulletin by the Board of Education of Enid in 1921.[2] Since the Enid plan of clasification has been operating for three years, it is now possible to present some statistical facts and some observations that should throw additional light upon the general question of classifying pupils according to ability. Some changes have been made in the original scheme of grouping. Two junior high schools have been added, changing the system from the eight-four to the six-three-three plan. All teachers[3] who are employed now are required to know how to test with the Stanford Revision of the Binet Test before entering upon their work.

A complete retesting has not been made, but all pupils new to the system are tested after trial in the schools. Retesting of some pupils has shown that those with low intelligence quotients usually test lower at each successive period. The group names

[1] By conservative estimate Enid has now a population of more than twenty thousand people. The nature of the community has not changed for several years, and similar results might be expected in any growing city.

[2] Price, E. D. *The Enid Plan of Classification of Pupils According to Mental Ability.* Enid, Oklahoma, Board of Education, 1921. 12 pp.

[3] That the teachers all favor the Enid Plan is evident from the answers to a questionnaire sent them by a student at the University of Oklahoma who was gathering data for a Master's thesis. Every teacher expressed unqualified approval of the plan, some going so far as to state that they would teach in Enid for a smaller salary rather than in a system not having a similar classification plan.

have been changed from Group I, Group II, and Group III, as they were first called, to Vocational Group, Rapid-Working Group, and Average Group.

In the present six-year elementary school the Average Group and Vocational Group progress by the scheme proposed in the tentative plan. In Table I the scheme is represented for the elementary school in terms of *Elapsed Time in School* and the grades completed. The average pupils cover six grades in six years, while the vocational pupils do four grades in the same

TABLE I. THE GRADE PROGRESS OF THE THREE GROUPS THROUGH THE ELEMENTARY SCHOOL AS SHOWN BY YEARS IN SCHOOL

	Elapsed Time in School—Years						
	1	2	3	4	5	6	
Rapid-Working Group	I–II–III		IV	V	VI	Promoted to J. H. S.	
Average Group	I	II	III	IV	V	VI	Promoted to J. H. S.
Vocational Group	I		II–III			IV	Promoted to J. H. S.

length of time. The "rapid workers" now do the first three years of the average course in two years and then join the "average workers" and keep step with them in grades IV, V, and VI. The one-year acceleration of these pupils makes the average fourth, fifth, and sixth grades difficult enough to keep them busy. These accelerated pupils, however, are always the leaders when grouped with older average pupils.

In similar manner the progress by grades in the junior and senior high schools for the three groups is shown in Table II. Progress is also expressed by the years of *Elapsed Time in School*. In the junior high school, grades VII, VIII, and IX, the rapid-working pupils may gain a half-year by completing three years' work in two and one-half years. In the senior high school the rapid-working pupils occupy the three years doing three years'

TABLE II. THE GRADE PROGRESS OF THE THREE GROUPS THROUGH THE JUNIOR AND SENIOR HIGH SCHOOL IS SHOWN BY YEARS IN SCHOOL

	Elapsed Time in School—Years										
	6	7	8		9		10		11		12
			First Half	Second Half	First Half	Second Half	First Half	Second Half	First Half		
Rapid-Working Group.....	VII, VIII, and one-fourth of IX		Last three-fourths of IX	First half of X	Second half of X	First half of XI	Second half of XI	First half of XII	Second half of XII	Graduated from S. H. S.	
Average Group.....	VI	VII	VIII		IX		X		XI		XII
Vocational Group.....	IV	V	VI		VII		VIII		IX		X

work. This group gains one year in the elementary school and half a year in the junior high school, a total gain of one and one-half years over the twelve years needed by the Average Group to complete the twelve grades of the public school. The Vocational Group finishes grade x in twelve years.

Of the total enrollment in the schools of Enid, 7 percent are classified in the Rapid-Working Group which includes pupils having I.Q.'s of 110 or above; 80 percent are in the Average Group; and 13 percent, those having I.Q.'s of 90 or less, are in the Vocational Group. The proportion allowed for the last group is

TABLE III. COMPARISON OF THE PERCENTS OF PUPILS PROMOTED, RETAINED, AND DROPPED DURING THE SCHOOL YEARS 1918-19 AND 1922-23

Grade	1918–19			1922–23		
	Percent Promoted	Percent Retained	Percent Dropped	Percent Promoted	Percent Retained	Percent Dropped
I....	76	15	7	88	10	2
II....	80	17	4	97	3	3
III....	84	13	3	94	3	0
IV....	82	15	3	94	5	1
V....	75	20	6	94	3	2
VI....	78	16	6	93	4	3
VII....	81	12	6	92	3	5
VIII....	76	8	16	92	4	2

slightly more than its actual enrollment, because it does not include all the children in town who have intelligence quotients below 90. A few have left school and in the lower levels are included the feeble-minded children who do not attend school. Because of the large enrollment of the average pupils, this group in each grade of the junior high school is divided into subgroups of 35 pupils classified by ranking the best thirty-five in the first subgroup, the next best in a second subgroup, and so on.

The curriculum of the elementary school is similar for the three groups. The course of study for the Rapid-Working Group in high school is more limited in electives than is that of the Average Group, but the former group is given more training in research. Vocational pupils are guided most carefully in their

subject-matter in order to give them such training as they can master. This article does not propose to tell in detail the subject-matter studied by each group, but it seems worth while to add that the care taken in suiting the work to the students seems to have resulted in success.

In Table III is given a comparison of the percents of pupils *Promoted*, *Retained*, and *Dropped* in the first eight grades during the school year 1918-19, before the classification scheme was in use and again in 1922-23 under the grouping regimen. The advantage of this method is shown by the consistent increase in the percent of promotions and a corresponding decrease in the percents retained or dropped.[4]

History of the Rapid-Working Group

In September, 1921, a group of pupils having high intelligence quotients and showing exceptional ability in their work entered grade IX of Enid High School. The number entering was fifty-six and the number finishing the year was fifty-four. The two who did not finish the first year left school to take positions as wage-earners. During the next summer vacation four of the group moved away so that only fifty entered in September, 1922. Forty-nine finished the year's work, one having dropped out during the year because of ill-health. Two pupils moved away during the following summer vacation, and forty-seven entered the following autumn. One pupil moved away during the year, leaving forty-six who graduated from the average four-year course at the end of their third year in school.

Another group of rapid workers entered in 1922 and a third group in 1923. Although they covered more subject-matter than the average pupils, their marks were in the upper ten percent of the school. One teacher of history made an investigation to see if the rapid workers were retaining the facts learned as well as the average workers who were going slower. At the middle of

[4] Such striking differences as are shown by Minnie Rutledge, University of Oklahoma, in her Master's thesis on the Enid School System leads one to believe that the advantage of giving all boys and girls a chance to succeed, as well as other advantages of the Enid Plan, is evident.

the year, the rapid workers had covered two-thirds of the text and the average pupils, one-half. An identical test, covering the first half of the text, was given to both groups. The rapid workers averaged over 5 points higher on the test, showing remarkably better insight in the interpretation of the facts of history. The investigation made with the history groups is typical of results all through the schools. The rapid workers who do the three primary grades in the first two years always make the highest marks in the fourth grade. Nor does covering more subject-matter in the junior high schools lower their rating.

History of the Vocational Group

In September, 1921, the first vocational group entered the high school which included at that time grades IX to XII. In this first group there were thirty pupils. Twenty-seven finished the first year; twenty, the second year; and seventeen, the third year. Of the thirteen who had left school, five took employment, as they were encouraged to do when a favorable position was offered. Four left town, and four were out but not regularly employed.

Another group of vocational pupils entered in September, 1922. There were thirty-seven of them. Thirty-one finished the first year and twenty-four, the second. Of the thirteen of this group out of school eight moved away, four are now regularly employed, and one is not.

Although vocational pupils are guided into suitable courses, they have some choice in their selection. The first year, for experiment, they were permitted to try some of the academic work. In 1921, members of the group were enrolled in eighty-nine vocational subjects and passed in sixty-nine of them. In 1922, they took seventy-eight vocational subjects and passed in sixty-eight. In 1921, they were permitted to enroll in twenty-nine academic subjects and passed in seven of them. In 1922, they enrolled in only three academic subjects and passed in but one. It is interesting to note that the vocational pupils soon learned to avoid the academic work and to select subjects suitable to their capacity.

In 1923, after a vocational group had been doing vocational work for two years in senior high school, some of the teachers felt that with their advanced age some of them might be able to do regular first-year work. Seven pupils were chosen who showed most ability and promise. The records of these five boys and two girls enrolled in first-year academic subjects follow (75 is the passing mark in the high school):

Paul—C. A., sixteen years, one month; I.Q., 89; took algebra, making a mark of 90; ancient history, 78; manual training, 84; English, 80.

Lyle—C. A., eighteen years, five months; I.Q., 78; took zoology, making a mark of 82; commercial law, 65; commercial English, 75.

Anna—C. A., sixteen years, nine months; I.Q., 72; took algebra, making a mark of 95; Bible, 80; ancient history, 80; English, 83.

Mary—C. A., sixteen years, eight months; I.Q., 72; took algebra, making a mark of 82; Bible, 87; ancient history, 83; English, 82.

Leonard—C. A., seventeen years, three months; I.Q., 88; took geometry, making a mark of 93; English, 80; medieval and modern history, 85.

William—C. A., eighteen years; I.Q., 89; took geometry, making a mark of 77; general science, 75; history, 86.

Burton—C. A., fifteen years, seven months; I.Q., 93; took algebra, making a mark of 73; manual training, 81; ancient history, 77; English, 70.

These marks show that some pupils of rather low intelligence after reaching a more mature age, can accomplish the regular work for beginning pupils in high school. Whether it is advantageous for them to take up the regular line of work is doubtful. Nevertheless, this arouses speculation, for the results show that pupils having lower intelligence quotients can be held in school until some of them can make passing marks in a regular course. This has not been true under the old plan of the "survival of the fittest."

Steps to Follow in Classifying Pupils According to Ability

The experience gained in thus classifying pupils according to ability should be of value to others who are contemplating a similar piece of work. The steps taken have proved so satisfactory and the results, including the good feeling between the parents and the school authorities, are so favorable that there

is no hesitation in describing the procedure for such benefit as others may derive from it.

1. The superintendent should study statistics and then make himself intelligently acquainted with scientific measurement.

2. With charts, tables, and whatever else is useful, the superintendent should give a series of lectures to his teachers until they understand his plans and are eager to assist in the undertaking.

3. The newspapers should be furnished with plain discussions of facts, showing that all pupils are not able to do work of equal amount or quality; emphasizing the handicap of "failure" in school and the danger of holding intelligent pupils back; and telling what other schools are doing to overcome these handicaps by intelligent classification.

4. The teachers should be taught to give the needed tests.

5. At the end of the first semester of a school year, there should be a check-up of the conditions in each room of the elementary and junior high schools. Since the senior high-school pupils are a selected group under traditional organization, it is probably well to omit testing them at this time. After it is found that many pupils will fail because of inability to accomplish the work, discuss the matter with the parents and give them the choice of having the pupils fail and repeat the work, or gain promotion at the end of the year in work they can accomplish. We have found it best always to consult with parents before placing a child either in the Vocational Group or in the Rapid-Working Group.

6. Reclassify pupils according to the teacher's judgment. The teacher in charge of each room should then test the pupils with individual mental tests, one each evening or morning. We have obtained the best results by having each teacher test her own pupils. Achievement tests should then be given and the results considered.

7. Make all necessary readjustments and proceed with the new schedule prepared for the system of schools.

8. Make the results of the classification according to ability frequently known through the local newspapers.

MINIMUM ESSENTIALS IN REPORTING DATA ON STANDARD TESTS

G. M. Ruch
State University of Iowa

With the increasing number of educational and mental tests, an already bewildering situation is daily becoming more aggravated. The writer refers particularly to the task of the superintendent, director of research, and the college teacher of tests and measurements, who is confronted with the problem of selecting and recommending the "best" test to use. Even if the cards were all on the table, the selection of the "best" test would present grave difficulties. These decisions must be made, and are being made daily, but it is an open question whether any living being possesses the exact knowledge required to make such decisions in anything approaching a scientific matter. Where does the difficulty lie?

The blame, to use a harsh term, lies primarily with the test authors, secondarily with the publishers of tests, and finally, to some extent, with the users of tests—the relative culpabilities showing wide individual differences within these groups.

The logic of the situation, viewed broadly, would seem to demand that two things be done at once: first, that a set of working *criteria* for test construction be established—and much progress has been made on this point; and second, that test authors, test publishers, test users, and test investigators adopt some fairly uniform practices in reporting on tests, at least to the extent of a few *minimum essentials*. Except in isolated cases, this has not been the regular practice. Although the second of these points is the primary consideration of this paper, a few statements about the first[1] will help to clarify matters.

[1]Kelley, T. L. "The Reliability of Test Scores," *Journal of Educational Research*, 3:370-79, May, 1921.
Monroe, W. S. *The Theory of Educational Measurements*. Boston, Houghton Mifflin Company, 1923. pp. 182-231.
McCall, W. A. *How to Measure in Education*. New York, Macmillan Company, 1922. pp. 195-410.

Criteria for Evaluating Tests and Measurements

Validity—the general worth or merit of the test.—A description of the validation of a test may well include such facts as the following:

1. The criterion against which the test was validated: analysis of courses of study, analysis of textbooks, recommendations of national educational committees, experimental studies (for example, word counts, analysis of social needs), age or grade rise in percent of successes, judgments of "competent" persons, correlation with an outside or independent criterion, or, merely, "armchair analysis" and "insight."

2. Statement of the exact details of all experimental work leading to the final forms of the test.

3. Statement of the diagnostic powers of the test, if any.

4. Statement of the exact field or range of educational or mental functions for which the test is claimed to be valid.

5. Statement whether the test is adequate for class measurement or pupil measurement. (This is largely a matter of reliability.)

6. Statement relative to equality of forms and how guaranteed. The same applies to equality of variabilities, namely, equal standard deviations for same groups.

7. Statement of the degree of correlation of the test with the criterion, or with school success, age, etc.; or statement of the agreement of test scores with the attainments of groups of individuals known to be widely spaced upon a scale of abilities.

8. Statement of the functions which the test can legitimately claim to serve, for example, prognosis, sectioning of classes, assignment of official grades, determination of passing or failing, efficiency of instruction.

These should be accompanied by summaries of the experimental evidence.

Reliability—the accuracy with which the test measures.—This is independent of its validity, although high validity demands high reliability. Such facts as the following are absolutely essential:

1. Reliability coefficients which are usually to be determined by correlation of similar forms. These are, however, practically meaningless unless accompanied by statements of: (1) the range of talent involved in the determination—best stated in terms of the standard deviation of the test scores; (2) the population involved in the determination of the *r*'s; (3) the age or grade groups involved and any evidence leading to judgments of the amounts of selection or systematic tendencies, or other factors militating against the representativeness of the sampling; (4) the order of giving the forms of the test; and (5) the mean scores on each form. These facts should be given in sufficient detail to permit a second investigator to reproduce the essential conditions of the experimentation at will.

2. Certain derived measures, which are relatively independent of: (1) the range of talent employed and (2) the arbitrariness of the test units.

Ease of administration and scoring.—The following facts might be given:

1. Degree of objectivity of scoring. This influences markedly the reliability, and hence, the validity of the test.

2. Time for giving. This is of minor importance without supplementation by other facts, the popular opinion to the contrary, notwithstanding. The proper criteria should be: *Validity per unit testing time, reliability per unit testing time,* or some similar point of view.

3. Time needed for scoring. This is best stated in terms of the average number of papers scored per unit of time. That this is of secondary importance is shown by the fact that comparatively few standard tests exceed the ordinary written examination in the time or labor of scoring.

4. Simplicity of directions for pupil and examiner.

Norms.—This should include:

1. Kinds of norms available (age, grade, percentile, mental age, T-scores, etc.)

2. Statement of derivation of norms. This should cover, specifically, facts like those listed under "Reliability" above. The important thing is the *representativeness* of the sampling,

not the size. Norms on one hundred thousand cases are not necessarily as accurate as those based on ten thousand, or even one thousand cases. *The validity of a norm is not determined by its probable error, but by the principles and laws of sampling observed or violated.* Norms, at best, are doubtful devices;[2] and blind faith in numbers approaches "hocus pocus" at times. The most important thing which a test can do is to *place the pupil in accurate rank positions along a scale of true ability.* The common practice of pooling test-score tabulations voluntarily returned to the author, without a scrupulous program for the elimination of the almost inevitable selection effects incident to this procedure, is particularly to be regretted.

Cost.—This has little or no theoretical interest but is a very practical consideration. The cost per pupil is practically valueless unless other facts are weighed. *Validity per unit cost* is the criterion to apply. A test may be a poor investment at one cent per pupil, while a second test would be cheap at ten cents per pupil. Costs of test vary with (1) the cost of experimental work (to the writer's knowledge, the variation on this point ranges from less than $100 to at least $10,000), (2) quality of printing, (3) length of test—a very important factor in validity and reliability, and (4) profits to publisher and author.

With respect to all of the before-mentioned desiderata, no exact procedure can or need be recommended. It is merely the "spirit of the game" that is fundamental. There is, however, one very important question to be asked, Where should the above data be published? The best place, theoretically and practically, is in the manual of directions accompanying the test. This can well be expanded to four, ten, or even one hundred pages. The next best procedure is probably that of publishing abstracts of the complete description of the test in the manual of directions, reserving the details for articles in the standard journals or for publication in a separate monograph. The important thing is

[2] *See* Chapman, J. C. "Some Elementary Statistical Considerations in Educational Measurements," *Journal of Educational Research*, 4:212-20, October, 1921 for an excellent treatment of this question, and *Manual of Directions Stanford Achievement Test*, Revised Edition; particularly the Appendix.

that full accounts be made accessible to the critical user or student of tests. To the user of tests should be extended the privilege of choice with open eyes, namely, with the "cards all on the table," to repeat an earlier statement.

If space permitted, the writer could at least entertain the reader by quoting numerous replies received from authors of secondary-school tests in response to an appeal for such data as have been outlined. Parenthetically it might be stated that fully 75 percent of test authors had made no systematic or critical study of their tests, and not a few did not comprehend the conventional test terminology. One responded to a question about the reliability of his test: "This is not an intelligence test." The correspondent must have been amazed at the writer's naïveté in expecting an educational test to possess reliability at all!

The Reporting of Reliabilities of Test Scores

The remainder of this discussion will be devoted to a single one of the criteria listed for the evaluation of tests, namely, *reliability*. Attention is directed to this topic partly because of its importance and partly because the data available on it are exceedingly meager. There has been a wide variety of practices in reporting test reliabilities—when, indeed, such have been reported at all—as follows:

1. The *reliability coefficient*[3] r_{12}
2. The *probable error of estimate*,[4] $P.E._{1.2} = 0.6745\,\sigma_1\sqrt{1 - r_{12}^2}$
3. The *probable error of measurement*,[4] $P.E._M = 0.6745\,\sigma\sqrt{1 - r_{12}}$, where $\sigma = \frac{\sigma_1 + \sigma_2}{2}$
4. The *index of reliability*,[5] $r_{11} = \sqrt{r_{12}}$
5. $\frac{P.E._M}{M}$, where the $P.E._M$ is given by formula 3 above,[5] and

[3] See any standard textbook on statistics.

[4] Monroe, W. S. *The Theory of Educational Measurements.* Boston, Houghton Mifflin Company, 1923. pp. 206 ff.

[5] Monroe, W. S. *A Critical Study of Certain Silent Reading Tests.* Urbana, University of Illinois, 1922. pp. 32 ff. (Bureau of Educational Research, College of Education, University of Illinois Bulletin Vol. XIX, Series, No. 8)

M is the mean of the distribution of scores, and, presumably,[5] equals, $\frac{M_1 + M_2}{2}$

6. The *probable error of estimate of a true score by means of a single score of the same function,*[6]
$P.E._{\infty_1} = 0.6745\ \sigma_1\ \sqrt{r_{12} - r_{12}^2}$

Examination of the six procedures just listed will show at once that *every one,* except the first, in part, *calls for two fundamental facts,* namely, (1) the reliability coefficient; and (2) the standard deviations of the two distributions.

To these must be added, at least (3) the population on which r_{12} is based in order to calculate the probable error of r.

The following are most desirable and probably should always be reported: (4) a verbal description of the kind of talent involved in the calculation of r, for example, age or grade group, kind of school, and possible selective factors militating against the representativeness of the sampling; (5) the mean scores of the two distributions leading to r; (6) the order of giving the tests[7] and the conditions under which the testing was carried out.

It is greatly to be regretted that one further recommendation is not always practicable, namely, (7) the publication of the scatter diagrams for all r's reported—this for its bearing on the lack of rectilinearity of the regressions and the possibility of faulty grouping. Both of these factors lower greatly the obtained r in comparison with the truth of the relationship.

Returning to our statements of the various methods of treating the reliability of test scores, it will be seen at once that our list of needed facts more than covers the needs of any or all of the six procedures, in fact, points (1) and (2) alone will suffice for the purely computational procedures. Granting this, the task of estimating the reliability of a test still would be a bit

[5] Kelley, T. L. *Statistical Method.* New York, Macmillan Company, 1923. p. 212 ff.

[7] Because the order of giving the tests must necessarily be different for the determination of reliability coefficients than for the investigation of equivalence of difficulties from form to form. For the first purpose, all pupils should take the tests in the *same* order, e.g., Form A followed by Form B; for the second purpose, one half of the group should take Form A first and one half Form B first. In the first case we want the practice effects to be systematic, in the second, they should tend to be neutralized.

of a "leap in the dark" without the subsidiary data outlined in points (2) to (7), inclusive.

It is rather beyond the scope of this paper, and beyond the ability of the writer as well, to demonstrate the absolute superiority of any one of the six proposals. However, certain of these have grave defects that must be made apparent in the interests of their abandonment for the purpose. The six methods will be commented on briefly in turn.

What Is the Best Method of Stating Test Reliabilities?

The first method, the reliability coefficient, can be held to be almost valueless, per se, for two reasons: (1) it is a function of the range of talent (σ) and hence has no general stability, and (2) alone, it tells nothing about the behavior of the individual score. For Examiner A to report r_{12} as 0.85 for Test X and Examiner B to report 0.64 for Test Y, *does not at all imply* that Test X is more reliable than Test Y. It would depend upon the range of talent employed. Assume the following facts for the case:

	Test X	Test Y
r_{12}	0.85	0.64
Group tested	500 pupils (Grades IV-XII)	500 pupils (Grade VI only)
Standard deviation	40.4	10.1

It may readily be shown that Test Y, if applied to the group to which Test X was given, would show a reliability coefficient in the neighborhood of 0.98.[8]

The needed measure of reliability must be independent, in large part, of the influence of range of talent. The reliability coefficient, as we have seen, is not. The practice of reporting r's unsupported by other facts should be discontinued.

The second proposed measure is the probable error of estimate,

---

[8] See Kelley, T. L. *Statistical Method.* New York, Macmillan Company. 1923. p. 222, formula 178;

$$\frac{\sigma_1}{\Sigma_1} = \frac{\sqrt{1-R}}{\sqrt{1-r}}, \text{ or, } \frac{10.1}{40.4} = \frac{1}{4} = \frac{\sqrt{1-R}}{\sqrt{1-.64}}. \quad R = 0.98-.$$

It is assumed that both tests are scaled to the same units.

$$P.E._{1.2} = 0.6745\ \sigma_1\ \sqrt{1 - r_{12}^2}.$$

This, however, does not serve the purpose. It helps little, or not at all, to obtain an estimate of the score in a second form of a test, when it is a simple matter to obtain a much better second score by actually giving the second test. An estimate of a *true* score is what is needed. The probable error of estimate does have a real value[9] as a measure of alienation from perfect prediction.

The third proposal (by Monroe, apparently) of a probable error of a test score is a development of the same formula. It has been shown that the correlation of obtained scores with true scores of the same function is

$$r_{1.\infty} = \sqrt{r_{12}}$$

hence the probable error of estimate of true scores from obtained scores is

$$P.E._{1.\infty} = 0.6745\ \sigma_1\ \sqrt{1 - r_{12}}$$

by substitution in the formula for the probable error of estimate. This, however, is not the needed formula, but the reverse, for example, the probable error of estimate of a true score from an obtained score. Such a formula is number 6 in the list, namely,

$$P.E._{\infty.1} = 0.6745\ \sigma_1\ \sqrt{r_{12} - r_{12}^2}.$$

The fourth proposal, namely, the ratio of $P.E._M$ to M, in the opinion of the writer, has no utility. Further comment is omitted here, because it has received an able criticism[10] since the first draft of this paper was written.

The fifth formula

$$r_{1\delta} = \sqrt{r_{12}}$$

is useful in evaluating certain correlation situations but would seem to have no particular reference to the problem of the reliability of test scores.

The last formula is probably the only one of the list which is entirely adequate to the problem at hand. It possesses the

[9] Especially in the form recommended by Kelley, i.e., $k = \sqrt{1 - r_{12}^2}$

[10] Franzen, F. R. "Statistical Issues," *Journal of Educational Psychology*, 15:367-82, September, 1924.

merits of being independent of the range of talent and of allowing for regression effects in test scores. The formula,

$$P.E._{\infty.1} = 0.6745\,\sigma_1\sqrt{r_{12} - r_{12}^2},$$

gives us the probable error of estimating *true scores* from obtained or fallible scores, when the true scores are estimated by the formula,[11]

$$\overline{X}_{\infty.1} = r_{12}X_1 + (1 - r_{12})M_1$$

$\overline{X}_{\infty.1}$ may be regarded as the best estimate possible of a true score, such as would be obtained by the average of an infinite number of obtained (X_1) scores. It is the estimated true score and its probable error that are needed and not the reverse as in the case of the third proposal above. Kelley suggests further the use of the

$$\frac{P.E._{\infty.1}}{\sigma_1}$$

ratio[12] as a measure of the improvement due to the use of the test.

In conclusion, it will readily be seen that the real need in reporting reliability data on test scores is the publication of a minimum of four things:

1. The reliability coefficient
2. The standard deviations of the two distributions
3. The population involved in the calculation of the r's
4. The means of the two distributions

These four data will permit the treatment of reliability by any of the methods proposed to date, and in addition, estimates of true scores, and prophecies of changes in reliability with changes in the range of talent. The estimate of regression effects is implied as another possible procedure; and, in cases of intercorrelations of test scores, the application of correction formulas for attenuation is made possible in evaluating true relationships

[11] Kelley, T. L. *Op. cit.*, p. 214, formula 168.

[12] Actually, $\frac{\sigma_{\infty.1}}{\sigma_1}$ appears in Kelley's recommendation, but either ratio leads to similar interpretations (*Loc. cit.*, p. 215). The σ_1 cancels in numerator and denominator, leaving the expression under the radical; namely, $\sqrt{r_{12} - r_{12}^2}$ as the important measure.

of the variables within the validity of the assumptions of such correction formulas.

Non-publication of such data as we have recommended is really a violation of the ethical codes of scientific procedures and not to be condoned by virtue of the fact that users of tests generally will not understand the technicalities. Rather, the teacher of tests and measurements should attempt to educate outgoing students to demand such confidences on the part of test authors. The alternative will certainly often be the refusal to recommend to school officials tests and scales upon which no critical facts are at hand. Such tests are not necessarily undependable, but the careful worker will not wish to assume responsibilities of proof which in all fairness rest upon the author of the test. The test buyer is surely entitled to the same protection as the buyer of food products, namely, *the true ingredients printed on the outside of each package.*

This statement, alone, is offered as a sufficient justification for presenting facts that are in no sense original with the writer.

SECTIONING JUNIOR HIGH-SCHOOL PUPILS BY TESTS AND SCHOOL MARKS[1]

FOWLER D. BROOKS
Johns Hopkins University

THE PROBLEM

This investigation was undertaken with the purpose of determining the value of fifth- and sixth-grade school marks and certain group tests of intelligence and achievement for dividing entering junior high-school classes into sections homogeneous in respect to scholarship; the accuracy of sectioning which may be secured from a few measures; and the procedures which may be used in school administration to secure the most accurate sectioning from the few measures employed.

Administratively it is impossible and educationally it is probably undesirable to give a great many tests for purposes of sectioning. Practical school administration necessitates such a relatively simple testing program that it may be an integral part of the administrative policy of the school system. The majority of junior high schools are not large enough to use many bases of classification in dividing first-year classes,[2] especially since the administrative difficulties of the junior high-school principal are complicated by semiannual promotions, by the two or more courses open to pupils sometime after the beginning of the first year, as well as by the frequent necessity of preserving a nearly equal proportion of boys and girls in each section on account of industrial work—shop, cooking, sewing, sheet metal—to say nothing of the task of maintaining a fairly constant teacher-pupil ratio.

Accordingly we have investigated the accuracy of sectioning which results from using a few measures, say two or three, of which one or two are available in the record system in the prin-

[1] Read before the Educational Research Association, Cincinnati, Ohio, February 24, 1925.

[2] If a class is divided into three sections—the top third, the middle third, and the lower third—upon each of n bases of classification, 3^n is the resulting number of sections.

cipal's office. Since the theoretically best combination of tests and school marks is too long and involved for wide use, I have suggested relatively simple ways of combining a few measures of high accuracy in sectioning.

Conclusions

I have set up a composite criterion of first-year junior high-school scholarship against which tests and school marks have been checked as bases of sectioning. The following conclusions seem warranted. The evidence supporting them is presented in the next section.

1. Relatively accurate sectioning is secured if the pupils entering the first-year class in junior high school are divided upon the basis of their sixth-grade school marks.

2. Sixth-grade marks give slightly more accurate sectioning than the average of fifth-and sixth-grade marks.

3. Dividing the first-year class of a certain junior high school into three sections on the basis of sixth-grade marks, at the same time putting approximately one-third of the boys and one-third of the girls in each section, and making two readjustments on account of differences in chronological age, gives 63 percent correct sectioning. By correct sectioning is meant that the pupil would be assigned to the same section by tests or school marks at the beginning of the first year in junior high school as he would be at the end of the year upon the basis of his scholastic success during that year.

4. Intelligence quotients from a group intelligence test give on the average, 55 percent correct sectioning, if a few readjustments are made on account of differences in chronological age;[3]

[3] After dividing pupils into three sections on the basis of group intelligence quotients the pupils of Section I were compared with those of Section II. If any pupils of Section II were more than six months older chronologically and had intelligence quotients not more than 2 points less than those of any pupils in Section I, such pupils were interchanged. And similarly for Sections II and III. This was done because of two or more pupils having the same intelligence quotients the older ones chronologically are, of course, the abler.

one of the achievement tests gives, on the average 47 percent correct sectioning.

5. Using sixth-grade marks and the mental ages or scores or intelligence quotients, from a group intelligence test such as the Haggerty Delta 2, and making three readjustments on account of differences in chronological age, we get 73 percent correct sectioning.

6. In sectioning on sixth-grade marks we find that two percent of the pupils really belonging in the highest section are placed in the lowest, while using sixth-grade marks and a group intelligence test reduces this sort of misplacement to one percent. In a three-fold grouping, placing in Section I a pupil who belongs in Section II may not be a very serious error; he may really be near the dividing line; but placing a pupil in Section III when he belongs in Section I is a serious error. As we have just seen, when we use sixth-grade marks and a group intelligence test, our data show this serious error in one percent of the cases.

7. Such serious errors of placement would have occurred, on the average, in 5 percent of our cases, if a group intelligence test had been used by itself, and in 12 percent of the cases if one of the five achievement tests had been used.[4]

8. If chronological age alone had been used—the youngest pupils being assigned to Section I, the next youngest to Section II, and the oldest to Section III—51 percent would have been correctly assigned; and 7 percent would have been seriously "mis-assigned."

9. Sixth-grade marks correlate 0.70 with the composite scholarship criterion for 93 pupils who had been in junior high school for one year. Intelligence quotients from group intelligence tests correlate from 0.50 to 0.72 with this criterion, the mean being 0.67. Achievement tests correlate from 0.38 to 0.62 with the scholarship criterion, the mean being 0.50.

10. Chronological age has a correlation of —0.49 with scholastic success in first year of junior high school.

11. The multiple correlations between two or more measures

[4] Our battery of achievement tests is not the best; it was selected for various purposes of which only one was predicting junior high-school success.

and our criterion of scholarship show the value of these measures in predicting scholastic success in junior high school. The multiple correlations between scholarship in first-year junior high school and the combination of sixth-grade marks and intelligence quotients from one group intelligence test range from 0.72 to 0.83, averaging 0.79; that is, using a group intelligence test with sixth-grade marks raises the correlation with junior high-school scholarship from two to thirteen points; on the average, the increase is 9.5 points (from .696 to .791). One of the achievement tests adds about 6 points, the multiple correlations ranging from 0.72 to 0.79, and averaging 0.755.

12. Chronological age adds two-thirds as much to the predictive value of sixth-grade marks as does a group intelligence test and, on the average, slightly more than does one of the five achievement tests used. The multiple correlation of sixth-grade marks and chronological age with the criterion is 0.76.

13. By adding a second group intelligence test or an achievement test such as the Holley Sentence Vocabulary Test to the combination of sixth-grade marks and a group intelligence test, we get a multiple correlation of 0.84 or 0.85 with first-year scholarship.

14. It is better to use sixth-grade marks and one group intelligence test, than to use two group intelligence tests. In the former case we get multiple correlations ranging from 0.72 to 0.83, averaging 0.79, whereas in the latter case they range from 0.74 to 0.78, averaging 0.757. Sixth-grade marks and chronological age, or one of the achievement tests, have as high predictive value as the intelligence quotients from two group intelligence tests.

The Evidence

Let us now examine the evidence upon which we have based the foregoing tentative conclusions.

Nine group intelligence tests and five achievement tests were given to an entering class in one of the Baltimore junior high schools.[5] The group intelligence tests in the order in which they

[5] The Robert Fulton School. I am indebted to Dr. Rozell Berryman, principal, for his cooperation in this and other investigations.

were given are: Miller, Form A; Otis Self-Administering; Illinois, Form A; Terman Group Test, Form A; Haggerty, Delta 2; National Intelligence Test, A1; Dearborn, Revised C and D; and Pintner Non-Language. The achievement tests, also in order of giving, are: Thorndike-McCall Reading, Form 3; Kelley-Trabue Language Completion Alpha; Holley Sentence Vocabulary; a group vocabulary test consisting of the first list of words from the Stanford Binet Test, two parts, 7 minutes to each part; and Woody-McCall Mixed Fundamentals in Arithmetic, Form A.

The elementary-school marks for the fifth and sixth grades and the chronological ages were taken from the pupil's permanent record card. The marks are upon a five-fold basis, E, G, M, P, and D—a mark for each half-year. They were given the following numerical values: E, 5; G, 4; M, 3; P, 2; and D, 1. Only the ninety-three who had complete records are here considered.

We need some standard or criterion of scholastic success in the first year of junior high school so that we may judge the value of tests and marks as bases of sectioning or prediction. The problem of a criterion is just as difficult as it is important. One difficulty with our criterion is that we do not know definitely or quantitatively the effect of clear-cut homogeneous grouping upon it. If a class is divided into homogeneous groups and school marks are then used, we must know about the standards which teachers have used in marking different sections. If the pupils in each section are marked according to some scheme based upon a priori assumptions of a predestined distribution of achievement, the distribution being taken the same for each section, then first-year marks will be of no value as a criterion. The correlation between marks and any truly differentiating test will of necessity be low or near zero. If, however, some reasonable standard is used, the criterion will be of value.

If we take as our sole criterion the results of achievement tests, we assume that the sampling in the tests really is so typical of the curriculum that we are safe in regarding relative

achievement in the tests as proportional to achievement in the whole curriculum. Many persons are not yet ready to make this assumption, especially in history, geography, and certain portions of English and mathematics in the first year of the junior high school.

If pupils are not grouped homogeneously, are the results the same as if they had been? This raises the fundamental question of the effect of homogeneous grouping upon achievement, a question which we cannot discuss at this time. Under certain circumstances achievement has been found to be more or less closely proportional to ability.

In view of these difficulties I have selected as the best available criterion of scholastic success—and probably as good a one as I could hope to secure—the following: (1) The average seventh-grade marks in English, mathematics, history, and geography given to pupils in fairly homogeneous groups by teachers using the same standards in marking different sections; (2) the educational ages from probably the best battery of achievement tests now available, the Stanford Achievement Tests. I have weighted the average marks 2 and the educational ages 1. These weights are purely arbitrary and were selected largely because there are important features of the first-year course of study in these four subjects, which are not included in the Stanford Achievement Tests.

The ninety-three pupils for whom I have complete records were in two sections during the first half-year in junior high school. They were scattered throughout four sections during the second half-year, grouped upon the basis of previous school marks and chronological age, in so far as administrative exigencies permitted.

For the purpose of this study I have divided these pupils into three sections upon the basis of the criterion of scholarship in the seventh grade, and have used this as correct sectioning. I next divided the pupils into three sections upon each of the fourteen group tests, sixth-grade marks, and the average of fifth- and sixth-grade marks. The comparison of the sectioning upon

each of these bases with that according to the criterion to find the amount of agreement and disagreement appears in Table I.

The high degree of accuracy of sectioning by sixth-grade marks is noteworthy; it is higher than that from any of the group intelligence tests or achievement tests which I used. It is as high as that from the Stanford-Binet Test. This is indeed surprising, especially when considered in connection with

TABLE I. PERCENT OF AGREEMENT OF SECTIONING BY TESTS AND SCHOOL MARKS WITH SECTIONING BY CRITERION OF SCHOLARSHIP IN FIRST YEAR OF JUNIOR HIGH SCHOOL. THREE-FOLD GROUPING

Basis of Sectioning	Correctly Sectioned	Displaced One Section	Displaced Two Sections
Miller (IQ)	54.8	40.9	4.3
Otis (IQ)	61.3	35.7	3.0
Illinois (IQ)	45.2	47.3	7.5
Terman (IQ)	63.4	32.3	4.3
Haggerty (IQ)	57.0	38.7	4.3
N. I. T. (IQ)	59.1	36.6	4.3
Dearborn C (IQ)	54.8	43.0	2.2
Dearborn D (IQ)	52.7	40.8	6.5
Pintner	49.5	39.7	10.8
Mean Group Intelligence Tests	55.3	39.4	5.2
Thorndike-McCall	50.5	40.9	8.6
Woody-McCall	52.7	34.4	12.9
Holley Vocabulary	45.2	43.0	11.8
Kelley-Trabue	39.8	45.1	15.1
Group Vocabulary*	47.3	40.9	11.8
Mean Achievement Tests	47.1	40.9	12.0
Stanford-Binet (IQ)	64.5	30.1	5.4
Chronological Age	51.6	40.9	7.5
Sixth-Grade Marks	63.4	34.4	2.2
Combination of Sixth-Grade Marks and Haggerty (IQ)	73.1	25.8	1.1

*Group vocabulary test made by using first list of Stanford-Binet vocabulary.

the interpretations often placed upon studies of the accuracy and reliability of teachers' marks. It indicates that accuracy of *term marks* should not be confused with that of the *marks assigned to a set of papers* in English, mathematics, or history; that the former may be and doubtless are much more accurate than the latter. From three to eight teachers assigned the junior high-school

marks. The pupils came from five elementary schools. It seems that when Baltimore sixth-grade teachers give a final term mark on this five-fold marking scheme they estimate not only present achievement, but also other qualities, abilities, or traits, which are predictive of work in the seventh grade. Sixth-grade marks were more closely related to the seventh-grade criterion of scholarship than to the seventh-grade marks alone.

Another way to determine the predictive value of these tests and school marks is to find the correlation between them and the criterion of seventh-grade scholarship. Then multiple correlations can be calculated to show the predictive value of different combinations of marks and tests. I have made these calculations. Sixth-grade marks and a group intelligence test correlate on the average .79 with scholarship in seventh grade. I found that Haggerty Delta 2, National A1, Dearborn C, and Dearborn D added most to the predictive value of sixth-grade marks, the multiple correlations being from .81 to .83. A second test such as National A1, Dearborn D, or Holley Sentence Vocabulary added to the combination of sixth-grade marks and Haggerty, raises the multiple correlation two or three points, that is, to .84 or .85. Chronological age and sixth-grade marks give a multiple correlation of .76 with the seventh-grade criterion.

Wherever I have combined scores from different tests I have followed the standard procedure of expressing an individual's score as a standard measure, that is, as a deviation from the mean of his group divided by the appropriate standard deviation. Standard measures have been weighted according to the proper partial regression coefficients.

On the whole, other studies corroborate the data presented in this paper. Upper-grade elementary-school marks have been found to correlate from .34 to .73 with high-school scholarship and group intelligence quotients from .28 to .69. School marks have generally been found to yield slightly higher coefficients. In one study in which a combination of eleven educational and psychological tests gave a higher prediction of scholarship in the first year of junior high school than the sum of all marks

received in the first six grades, we must remember that fifth- or sixth-grade marks probably have much higher predictive value than those assigned four or five years before entrance to junior high school.

Combining School Marks and Test Results for Sectioning

Of course, the standard procedure in combining sixth-grade marks and test results is to take the average or mean of each set of measures, find the standard deviation of each set, take each individual's deviation in each measure from the mean of his group and divide this deviation by the standard deviation. These standard measures can now be combined by any desired scheme of weighting, or two multipliers may be chosen such that they will make the standard deviations equal. For example, if the standard deviation of the first set of scores is 4.5 and of the second set 3.0, the multipliers would be 2 and 3—the individual deviations in the first set of measures would be multiplied by 2 and those in the second set by 3, and then weighted and combined. Weighting is difficult. Our own data indicate that, on the average, sixth-grade marks and intelligence quotients from a group intelligence test should each be weighted 1, but our results are only tentative.

One of the following ways of combining marks and test results, while statistically less orthodox, is much simpler and nearly as accurate:

1. Rank the pupils in each measure and then combine the ranks by adding.

2. Assign decile ranks to pupils in each test or in school marks; that is, give a rank of ten to the 10 percent of pupils having the highest marks or scores or intelligence quotients, a rank of nine to the next 10 percent, etc. Do this for each test used and for school marks. Each pupil's decile ranks can then be combined readily. If school marks and test results are given equal weight, ranks or decile ranks can be added directly; if unequal weight, each pupil's rank in marks can be multiplied by the proper weight, and similarity for ranks in a test, and the weighted ranks then added.

Chronological age can be utilized by placing the older pupils of a given standing on combined marks and test scores in a slower section, and the younger pupils of this same standing in a brighter section. If intelligence quotients are used instead of scores or mental ages, the older pupils of a given intelligence quotient should be placed in a brighter or more advanced section than are the younger pupils of this same intelligence quotient.

In conclusion: The "fly in the ointment" is that we do not know well enough the predictive value of tests and marks in varying situations. We have data on 93 pupils in a junior high school, on 108 pupils in the same school, and on still another class of 303 pupils.[6] We also have some evidence from a class of 220 pupils in another school showing the high predictive value of sixth-grade marks and group intelligence tests. Further investigation is needed. Bureaus of research or junior high-school principals and their staffs undoubtedly would find it profitable to discover the predictive value for their own schools of tests and elementary-school marks. If several classes are studied the probable constants of certain predictive factors, such as school marks, group- or achievement-test results, and chronological age, may be determined with a reasonable degree of accuracy.

As an aftermath of the hue and cry about the unreliability and inaccuracy of teachers' marks many administrators have come to regard them as unclean, as necessary at times for certain purposes but to be used as little as possible. While term marks[7] are not as accurate or reliable as we might wish, yet they undoubtedly have considerable value, especially fifth- and sixth-grade marks, in prophesying scholastic success in the first year of the junior high school. Even the maker of intelligence tests usually falls back upon teachers' estimates to help validate his tests.

Combining the results of a group intelligence test with sixth-grade marks, and taking into consideration chronological age.

[6] Correlation between sixth-grade marks and seventh-B marks = .804 ($n = 108$); correlation between sixth-grade marks and seventh-B marks = .764 ($n = 303$).

[7] We are not discussing marks given by teachers to a set of papers. The unreliability of such marks has been wrongly confused with that of term marks which are based on daily grades, quizzes, examinations, attitude, effort, etc.

we have a fairly accurate basis for sectioning entering junior high-school pupils, a basis more accurate than any one of the three elements of which it is composed. Undoubtedly other factors are present, but they have not been adequately analyzed or estimated. Certain volitional, emotional, or interest factors may eventually be found to add much to a composite prognosis measure. Absolutely perfect predictive measures may never be found just because growth or development may of itself introduce a certain amount of change or variation which will continue to elude accurate estimate.

NOTES ON THE CURRICULUM IN COLONIAL AMERICA

ROBERT FRANCIS SEYBOLT
University of Illinois
(Continued from November)

Dr. Seybolt wrote the following paragraph with the intention of having it appear at the beginning of the first installment of this article. Unfortunately it was received too late to be used as he intended.—*Editor.*

The little studies that are grouped under this title are but notes on the teaching of certain subjects in the colonial period. The first two are limited to the Province of New York; the others refer to the colonies in general. Their purpose is to present materials not used by other writers, and to call attention to errors that appear in current texts. Obviously, they are not complete: intensive research will yield much that is merely suggested here.

BOOKKEEPING

In *A Contribution to the History of Commercial Education*, Mr. E. M. Barber takes occasion to criticize the following statements by "a recent writer on the history of early business education in America":

> The science of double-entry bookkeeping as known and used to-day was not known in 1840. An earnest young man then living in Philadelphia wishing to learn bookkeeping applied to many merchants, but none would teach him or accept him as a bookkeeper without experience. He went to New York and found no one there to employ him or teach him bookkeeping, as each merchant had his own way of keeping accounts.
>
> Bookkeeping was not then, as it is now, a system. It had no well-defined rules and was not taught save under the eye of business men themselves.
>
> About 1845, the first business school in New York was started by Benjamin Foster, a native of England, and the instruction given was in penmanship and bookkeeping.

Mr. Barber proceeds, in rather direct fashion, to expose the writer of these irritating paragraphs:

> The last date quoted, it will be observed, is 1845. Now, these statements come into collision with the truth at so many points as to remind one of the famous Academician who defined a crab

to be "a small red fish that walks backwards," a definition which proved to be true excepting in three particulars: it was not red; it was not a fish; and it did not walk backwards.

Prompted by curiosity, some weeks ago, I determined to ascertain, if possible, why the City of New York, foremost in commerce and finance, should have been ignorant of double-entry bookkeeping and deprived of commercial teachers and schools, prior to the year 1845.[1]

After some research, Mr. Barber finds, in the New York City "directory for 1818 . . . . the name of James Bennett, Accountant," who began his "Public Lectures on Bookkeeping," either in that year or in 1824.[2] But there was another Bennett in New York at this time, one James Gordon Bennett, who advertised, in 1824, "an English classical and Mathematical school for the instruction of young gentlemen, intended for mercantile pursuits."[3] This school, "if ever actually formed, had a brief existence." In 1825, James Gordon Bennett "returned to newspaper work."

> If New York's claim to priority were otherwise jeopardized, I should contend that this was the first commercial school founded in America. Curiously, that honor belongs, I believe, to "another man by the same name." James Gordon Bennett may have doubted his ability to withstand the James Bennett competition, and, probably, this had something to do with the abandonment of his proposed permanent commercial school."[4]

C. A. Herrick, who had examined Mr. Barber's "careful study," says:

> The probable facts are that the spirit of the technical business school of this country can be traced back to R. M. Bartlett, and that there has been a continuous development of this institution from his crude beginnings. It seems probable that Bartlett began his first school for the teaching of penmanship, bookkeeping, and commercial arithmetic at Philadelphia in 1834.[5]

---

[1] Barber, E. M. "A Contribution to the History of Commercial Education." (An address delivered at the Annual Meeting of the Eastern Commercial Teachers Association, held at Brooklyn, New York, April, 1903) p. 1.

[2] *Ibid.*, p. 5.

[3] *Ibid.*, p. 7.

[4] *Ibid.*, p. 8.

[5] Herrick, C. A. *Meaning and Practice of Commercial Education.* New York, Macmillan Company, 1904. p. 179.

"To whom belongs the honor of the first venture in commercial education is a matter of some dispute. It has frequently been attributed to . . . . R. M. Bartlett . . . . By others the honor is claimed for James Bennett." *Cyclopedia of Education.* Paul Monroe, editor. New York, Macmillan Company, 1911. Vol. II, p. 144.

Both Barber and Herrick have overlooked certain important sources of information on this topic. The colonial newspapers record the existence of schools in which commercial training was given for a full century before the ventures of Bennett and Bartlett.

Strange as it may seem, bookkeeping was taught in Boston, as early as 1709:

> Reading, Writing, Arithmetick, Merchants Accompts, Geometry, Trigonometry, Plain and Sphaerical, Dyalling, Gauging, Astronomy, and Navigation are Taught And Bonds, Bills, Indentures, Charter-parties, &c. are drawn; and Youth Boarded in Cross-street, Boston. By John Green.[6]

There it is: "merchants' accompts." In 1718, "Mr. Browne Tymms Living at Mr. Edward Oakes Shopkeeper in Newbury-Street at the South End of Boston" taught "Young Men Arithmetick and Merchants Accounts."[7]

Although similar New York records are not available for these years, it may be assumed that bookkeeping was taught there as early as in Boston. Both cities were important commercial centers before the close of the seventeenth century, and there must have been some response to the demand for business training. George Brownell, of New York City, in 1731, offered "merchants' accounts;"[8] and in 1734, Alexander Malcolm announced instruction in "merchant's book-keeping."[9]

In Philadelphia, 1733, Andrew Lamb advertised that he continued "to teach Merchants Accompts, after the Italian Manner

---

[6] *Boston News Letters*, March 21-28, 1709.

Boston Gazette, August 31-September 7, 1724. Samuel Grainger: "Writing, Accompts and the Mathematicks." Also September 4-11, 11-18, 1727. Samuel Grainger: "Writing, Arithmetick, Book-keeping."

[7] *Boston News Letter*, February 24-March 3, 1718. Also March 5-12, 12-19, 1730. Charles Lewis: "Writing, Arithmetick, Merchants Accompts, Foreign Exchanges, either in French or in English."

[8] *New York Gazette*, August 31-September 7, 1731.

New York Weekly Journal, March 24, 31, and April 7, 1740. James Foddey: "Accompts."

[9] *New York Gazette*, December 30-January 7-14, and 14-21, 1734.

New York Weekly Journal, April 6, 13, 27, and later issues in 1741. John Campbell: "Merchants Accompts."

by double Entry, Dr. and Cr. by the best Method."[10] The subject was taught in Charleston, South Carolina, 1744, by Stephen Hartley,[11] George Brownell, and John Pratt;[12] and in Chester Town, Maryland, 1745, by Charles Peale.[13] The Italian method, or double-entry bookkeeping, appears in many schools in eighteenth-century America.[14]

The usual commercial course of the period included "mercantile arithmetic," bookkeeping, and penmanship. For those who were preparing to enter the employ of merchants engaged in foreign trade, the schools offered instruction in French, Italian, Portuguese, Spanish, and foreign exchange.[15]

Shorthand

Professor Andrews, in a chapter on "The Intellectual Life" in the American colonies, says:

> *Whether shorthand was anywhere taught is doubtful and highly improbable,* yet from Henry Wolcott, Jr., of Windsor and Roger Williams of Rhode Island to Jonathan Boucher of Virginia and Maryland there were those who were familiar with it, and occasional references to writings in "characters" would point in the same direction.[16]

Frequent references to shorthand by men like those just mentioned are evidence of its use by ministers and writers of the period. Many others learned the "art of character-writing"[17] for the purpose of taking down and preserving sermons.

---

[10] *American Weekly Mercury*, September 6-13, 13-20, December 14-21, 1733, and in many issues during 1734 and 1735 appeared the notice of Theophilus Grew: "Merchants Accompts."

Pennsylvania Gazette, August 18-25, 1737. William Robins: "Book-Keeping."

[11] *South Carolina Gazette*, June 4, 11, and 20, 1744.

[12] *South Carolina Gazette*, September 3, 10, and 17, 1744. George Brownell and John Pratt: "Merchants Accompts, in the true Italian Method of Double Entry, by Debtor and Creditor."

[13] *Pennsylvania Gazette*, March 12, 19, 26, and April 2, 1745.

[14] Seybolt, R. F. *Source Studies in American Colonial Education*. Urbana, Illinois, University of Illinois, 1925. Chap. III.

[15] *Ibid.*, Chaps. I, II, III.

[16] Andrews, C. M. *Colonial Folkways*. New Haven, Yale University Press, 1921. p. 142-43. Italics mine.

[17] Also called the "Art of Characterie," "the Art of Swift Writing," "Brachygraphy," "Cryptography," "Palygraphy," "Tachygraphy," "Thoographia," during the seventeenth and eighteenth centuries.

Another writer states that "shorthand writing was taught from about the middle of the nineteenth century; but it was taught as a science or for professional rather than commercial purposes."[18]

It is impossible to establish the earliest date for the teaching of shorthand in this country. Records at hand indicate that it was offered in Boston as early as 1728:

> Caleb Philipps Teacher of the New Method of Short Hand, is remov'd to the north door of the Town House in King-street. As this way of Joyning 3, 4, 5, &c. words in one in every Sentence by the Moods, Tenses, Persons, and Verb do's not in the least spoil the Long Hand, so it is not anything like the Marks for Sentences in the Printed Character Books being all wrote according to the Letter, and a few Plain and Easy Rules.
>
> N. B. Any Persons in the Country desirous to Learn this Art, may by having the several Lessons sent Weekly to them, be as perfectly instructed as those that live in Boston.[19]

The next announcement of Caleb Philipps' school was published by a group of prominent citizens of Boston,[20] who took that occasion "to recommend his Art":

> WHEREAS Mr. Caleb Philipps arriv'd here last Fall from England, a Master of the celebrated Mr. Weston's Short-Hand; which has the fame of being the quickest, most regular, and easiest learnt of any; and having seen to our Satisfaction his Dexterity in writing it, as also something of his Aptness to teach it by the Improvement which one in particular has already

---

[18] Herrick, *op. cit.*, p. 192. "By 1849 shorthand had become popular in America." (p. 214)

[19] *Boston Gazette*, March 18-25, 1728.

[20] *Boston Gazette*, April 1-8, 1728.

Caleb Philipps' sponsors were the most prominent clergymen of Boston, in 1728.

Benjamin Colman (Harvard, 1692; A.M.; S.T.D., Glasgow, 1731; Fellow of Harvard College; died 1747) was pastor of the Brattle Street Church. In 1724, he was elected President of Harvard, but declined.

Joseph Sewall (Harvard, 1707; A.M.; S.T.D., Glasgow, 1731; Fellow of Harvard College; died 1769) was elected President of Harvard in 1724, but declined.

Thomas Prince (Harvard, 1707; A.M.; died 1758) was pastor of the Old South Church.

William Cooper (Harvard, 1712; A.M.; died 1743) was elected President of Harvard in 1737, but declined "this honor and trust."

Thomas Foxcroft (Harvard, 1714; A.M.; died 1769) served on committees of the Harvard Board of Overseers.

Joshua Gee (Harvard, 1717; A.M.; Librarian of Harvard College; died 1748) succeeded Cotton Mather as pastor of the North Church.

made under him; and being, by what we are inform'd, of a blameless and laudable Conversation—We Judge it may be for the publick Good to recommend his Art to all that are like to Have Occasion for writing much; especially to Scholars, and those that would preserve the Sermons they hear in public for their further and lasting Use.

By the Prints from London it appears that this Method has so recommended itself, that Mr. Weston has obtained His Majesty's Royal Authority for the sole printing and publishing his Books for the space of fourteen Years; and we are also credibly inform'd that many of the Nobility, Gentry, Ministry, Etc. have taken the Pains to learn it; and we are persuaded it would be of great Advantage, if it should be commonly receiv'd in this improving Country.

Boston
April 2, 1728

BENJAMIN COLMAN
JOSEPH SEWALL
THOMAS PRINCE
WILLIAM COOPER
THOMAS FOXCROFT
JOSHUA GEE

Philipps was quite up-to-date in his method of instruction. Weston's text, *Stenography Completed*, was published in 1727, the year in which Philipps appeared in Boston.

Stenography Completed, or the art of Shorthand brought to perfection; being the most easy, exact, lineal, speedy, and legible method extant; whereby can be joined in every sentence at least two, three, four, five, six, seven, or more words together in one, without taking off the pen, in the twinkling of an eye; and that by the signs of the English moods, tenses, persons, participles, etc., never before invented. By this new method any who can but tolerably write their name in round-hand may, with ease, (by this book alone, without any teacher) take down from the speaker's mouth any sermon, speech, trial, play, etc., word by word; though they know nothing of Latin; and may likewise read another's writing distinctly, be it ever so long after it is written. To perform this by any other shorthand extant is utterly impossible, as is evident from the books themselves. The nature, use, and excellency hereof, are more fully contained in the preface. Composed by James Weston, the only author and professor of this new method. London. Printed for the author, and sold by him at the "Hand and Pen," over against

> Norfolk street, in the Strand; where he continues to teach this new method expeditiously.[21]

At least thirty different systems of shorthand had been published in England by 1727.

The art was not unknown in Philadelphia in 1729. Samuel Keimer, printer and book-seller, finding himself in financial difficulties, offered to give away shorthand texts as premiums, in order to increase his sales:

> The said Keimer (Samuel Keimer, printer of the *Pennsylvania Gazette*) has re-printed a fresh Impression of Short-Hand Books, to give to every one that shall lay out but Three Shillings in Bonds, Bills, Indentures, Primmers, Accidences, or other useful Books, he selling cheaper than ordinary in order to the more speedy enabling him to do Justice to his Creditors.[22]

Following the transfer of his business to David Harry, Keimer published many notices requesting his debtors to pay up.

In New York City, 1751, "a new invented Short-Hand" was taught with reading, writing, arithmetic, book-keeping, and other subjects, in a school conducted by Benjamin Leigh and Garrat Noel.[23] A "new and concise Method of Writing Short Hand" was offered, in New London, Connecticut, in 1777, by "A Young Man that can be well recommended."[24]

Clerks of courts and town meetings, and secretaries to certain officials, as well as "Scholars, and those that would preserve the Sermons they hear," made use of shorthand. Although it had not come into general use in the business life of the colonies, there seems to have been a demand for instruction in "swift Writing" in the larger cities of eighteenth-century America.

---

[21] Pitman, Isaac. *A History of Shorthand.* 3rd edition. London, Sir Isaac Pittman & Sons, 1891, 19-20.

[22] *Pennsylvania Gazette*, Eighteenth of the Seventh Month, 1729.

[23] *New York Gazette* Revived in the *Weekly Post Boy*, January 21, 28, 1751. Also, February 25 and March 4, 1751. An unnamed master "teacheth a new-invented Short Hand."

[24] *Connecticut Gazette and Universal Intelligencer*, August 29, September 12, 1777.

School-Equipment

According to Wickersham, "slates and pencils did not come into use until after the Revolutionary War."[25] This observation is repeated by Cajori.[26] Clifton Johnson says: "Slates did not come into general use until about 1820, and lead pencils not for a good many years after that."[27] This may have been the source on which Cubberley relied for his statement that "slates were not introduced until about 1820, and pencils and steel pens did not come into use until much later."[28]

Such statements make one wonder why slates and pencils were advertised by various booksellers of the eighteenth century. Did no one purchase or use them? Joshua Blanchard, of Boston, sold "Slates," and "black lead Pencils," in 1743.[29] The advertisements refer to "Pencils," "Pencils with or without Steel Cases," "Best black lead pencils with silver cases," "Middleton's very best Lead Pencils."[30]

One author informs us that "The colonial schools had no blackboards and no maps, but once in a while a schoolhouse in the more flourishing communities would possess a globe;"[31] another, that "During the first half of the century after the War of the Revolution . . . . Charts, maps, and globes had not come

[25] Wickersham, J. P. *A History of Education in Pennsylvania.* Lancaster, Pennsylvania, Inquirer Publishing Company, 1886. p. 192.

[26] Cajori, F. *The Teaching and History of Mathematics in the United States.* Washington, D. C., Superintendent of Documents, 1890. (U. S. Bureau of Education, Circular of Information, 1890, No. 3), p. 10.

[27] Johnson, C. *Old-Time Schools and School-Books.* New York, Macmillan Company, 1904, p. 38. *Cyclopedia of Education.* Paul Monroe, editor, New York, Macmillan Company, V, 341: "Slates were rather slowly introduced into our schools and rarely found before the close of the Revolutionary War."

[28] Cubberley, E. P. *Public Education in the United States.* New York, Houghton Mifflin Company, 1919. p. 36.

——— *Readings in the History of Education.* New York, Houghton Mifflin Company, 1920. p. 369 quotes Noah Webster: "Before the Revolution, and for some years after, no slates were used in common schools."

[29] *Boston News Letter,* December 29, 1743.

New York Gazette or Weekly Post Boy, January 12, February 16, 1764. William Cockburn sold slates.

[30] *New York Mercury,* April 9, 1759; June 23, 1760; September 1, 1777; September 20, 1779; October 30, 1780.

[31] Johnson, *op. cit.,* p. 37.

Cajori, *op. cit.,* p. 10: "School appliances in those days were wholly wanting (excepting the ferule and birch rods)."

into general use."[32] Although blackboards are not mentioned in the sources examined, it is probable that they were in use here and there. Christopher Dock's blackboard could not have been the only one in eighteenth-century America.

The records reveal a different situation, however, with respect to the use of maps and globes. Theophilus Grew, of Philadelphia, 1735, taught "the Use of Globes, Maps."[33] In 1745, Charles Peale, "At Kent County School, near Chester Town, Maryland," gave instruction in "the Use of the Globes, by the largest and most accurate Pair in America."[34] John Nathan Hutchins, of New York City, 1763, announced that he had "a good Sett" of globes, both "Celestial and Terrestrial."[35] Maps, and "the making of Maps" are mentioned frequently in the school announcements of the period.[36] In the seaport cities, all the "better" schools, in which navigation and geography were offered, were equipped with globes, maps, and charts, and the students were taught to use them.

[32] Swett, J. *American Public Schools.* New York, American Book Company, 1900. pp. 118-19.

Cubberley. *Public Education in the United States.* p. 36: "There were no blackboards or maps."

[33] *Pennsylvania Gazette,* October 16-23, 1735. Also during November and December issues of 1743.

Charles Fortesque in Philadelphia advertises that he will teach "Use of the Globes."

James Cosgrove of Philadelphia makes a similar announcement in the issues for September 8, December 8, 22, 1757.

[34] *Pennsylvania Gazette,* March 12, 19, 26, April 2, 1745.

Boston News Letter, March 3, 10, 1743. Nathan Prince, Boston: "The Use of the Globes."

Independent Advertiser, April 17, 24, May 1, 1749. John Leach, Boston: "The Use of the Globes and Charts."

[35] *New York Mercury,* April 25, May 2, 1763.

New York Gazette. Revived in the *Weekly Post Boy,* January 21, 28, 1751. B. Leigh and G. Noel, New York City: "the Use of the Globes."

New York Gazette or Weekly Post Boy, January 12, February 16, 1764. Wm. Cockburn.

[36] *Pennsylvania Gazette,* September 12, 1771. Maguire and Power, Philadelphia: "the use of globes and maps, and how to make maps."

New York Gazette or Weekly Post Boy, May 12, 19, 26, 1755. Robert Loeth, New York City: "the Construction of the Plain and Mercator's Chart."

New York Mercury, May 6, 13, 20, September 30, October 7, 1765. Thomas Carroll, New York City.

New York Gazette and Weekly Mercury, October 28, November 4, 11, 1771. James Conn, Elizabethtown, New Jersey: "the most useful and most difficult Part of Geography, viz. Drawing Maps and Charts, whether Plain, Mercator, Spherical or Conical; together with their Explanation, and the Reason why each Kind are drawn in the Manner they are."

SOME EXPERIMENTAL STUDIES OF THE RESULTS OF COLLEGE NOTE-TAKING

C. C. CRAWFORD
University of Idaho

In an article in the November issue of this journal,[1] the writer summarized certain correlations which he had found to exist between lecture notes and the subsequent quiz grades of college students. The data used in the study did not permit an analysis of the effects of note-taking under experimental conditions. In this article an attempt is made to analyze further the different factors involved in note-taking. A series of experiments will be reported which include groups of students taught by three different instructors and which involve three quite different types of classroom procedure, namely, the formal lecture, the class discussion based on library reading, and the recitation upon the textbook. Furthermore, the conditions under which the notes were later reviewed varied considerably in the different experiments. Although this study is fairly comprehensive, it by no means exhausts the list of circumstances under which students take notes and use them. The results reported must, therefore, be judged in the light of the conditions under which they are procured, and the reader must exercise caution in generalizing upon these findings and in applying the conclusions here made to situations radically different from those which prevailed during the study.

The problem considered is two-fold: First, judged by the results of immediate recall, is it more effective to take notes as a lecture is being delivered or to listen without taking notes? Second, judged by the amount of the information retained, several days or weeks after a lecture, is it more effective to take notes at the time and review them later, or to take no notes and to depend upon memory or other sources of aid in reviewing? Another problem closely allied to our study is the investigation of

[1] Crawford, C. C. "The Correlation between College Lecture Notes and Quiz Papers," *Journal of Educational Research*, 12:282-91, November, 1925.

qualitative differences between the results of note-taking and of listening. It is reasonable to expect that two radically different methods of procedure during the class period may produce a different kind, as well as a different amount, of mastery. We have, therefore, used the true-false test and the traditional quiz side by side in several experiments to detect, if possible, any such qualitative differences. Two types of experiments were used, one in which the results were measured immediately after the lecture period and the other in which the results were measured after an interval sufficiently long to permit a considerable amount of forgetting and to make it necessary to use some kind of review. The first three of the seven experiments reported here were to measure the effect of note-taking and the last four to measure the net results of note-taking plus review. Each was a rotation experiment arranged according to the method described by McCall.[2]

PROCEDURE OF EXPERIMENTS MEASURING THE IMMEDIATE RESULTS OF NOTE-TAKING

Experiment I.—A class of 38 students, chiefly Sophomores, was divided into two groups, A and B. During the first lecture Group A was asked to take notes while Group B listened; a twenty-minute quiz of the traditional type followed. The next day the program was reversed, Group A listened while Group B took notes, and a quiz, similar to that of the preceding day, was given. Each student received a code number to use instead of his name, so that each paper was graded without the marker knowing the group to which the writer belonged. Similar precautions were employed in all the experiments.

Experiment II.—The procedure in this experiment was essentially the same as that in Experiment I, but in addition to the quiz a true-false test was given each day. Representatives from the four college classes were enrolled among the 129 students in this group.

[2] McCall, William A. *How to Experiment in Education.* New York, Macmillan Company, 1923. pp. 19-32.

Experiment III.—The students and teacher of Experiment II were used again in this section. The two procedures were similar, but in this experiment the testing was more comprehensive. Each true-false test in Experiment II contained ten questions and the general quiz period was eight minutes, but in Experiment III the true-false test included fifteen questions and the quiz period was lengthened to fifteen minutes.

PROCEDURE OF EXPERIMENTS MEASURING THE REVIEW VALUE OF NOTES

Experiment IV.—This was a rotation experiment involving 34 students, chiefly Juniors. After each of two lectures the students' notes were collected and held for a week. Then notes on one lecture were returned to Group A and notes on the other to Group B. Five minutes were allowed for review of the notes, after which both groups were given a written quiz of one hour covering both the lectures. Thus each group was questioned on one lecture which had been reviewed and on one which had not.

Experiment V.—This phase of the study was but a continuation of Experiments II and III, and the same students were used as subjects. At the close of a series of four lectures, the notes, which each student had taken, were returned to him, but since the groups had taken notes on alternate days, each group had notes on the two lectures to which the other group had but listened. A period of ten minutes was allowed for reviewing the notes. This was followed by a test in two parts; Part I included sixty true-false questions, grouped according to the lectures on which they were based, and Part II consisted of four general questions, one on each of the four lectures. The two experimental factors compared, therefore, were (1) taking notes and reviewing them and (2) listening to the lectures and trying to retain them without notes and without review.

Experiment VI.—Another factor appears in this experiment, for the class periods were devoted to discussions based on the library reading done by the students. The class consisted of 35 Sophomores, but because of absences only twenty cases were

available at the close of the experiment. The class was divided into two groups, which took notes on alternate days. The experiment covered six class days during a period of two weeks, followed by two days for testing results. The first test day was devoted to a general quiz of one hour, consisting of one comprehensive question on each of the six class discussions. On the second test day was given a true-false test consisting of ten questions on each of the six class discussions.

Experiment VII.—This experiment included 55 students, chiefly Seniors, and was very much like Experiment VI. The principal ways in which it differed from the former were: (1) the class periods were devoted to recitations on the textbooks and not to class discussions of library readings; (2) no true-false test was given; (3) there were eight days in the series instead of six; and (4) one day was devoted to a review directed by the instructor before the final quiz was given.

Results

Limitations of space forbid the inclusion of the detailed scores of the subjects tested, but the means and standard deviations of the different sets of scores are given together with the net differences between the results with and without notes. The reliability of each net difference is shown by its "experimental coefficient."[3]

The results of note-taking and of the use of notes as measured by the general quiz of the traditional type are given in Table I. In each of the seven experiments there is an advantage in favor of notes. This advantage, small in Experiments I and VII, is large enough to be of considerable importance in the others. The results of the different experiments can best be determined by the comparison of their respective experimental coefficients. This method of comparison is not altogether valid because the number of students in the different groups varied. Experiments II, III, and V involved over 120 subjects each, whereas the others involved from 20 to 55 subjects. Since the

[3] McCall, *op. cit.*, pp. 154-58.

TABLE I. RESULTS OF EXPERIMENTS BASED ON SCORES IN THE GENERAL QUIZ

TABLE I. RESULTS OF EXPERIMENTS BASED ON SCORES IN THE GENERAL QUIZ

Experiment	Kind of Value Measured	Net Score with Notes[a]	Net Score without Notes	Difference in Favor of Notes	Standard Deviation of the Difference (SDD)	Experimental Coefficient (Ec)	Approximate Chances That the Real Difference Is above Zero
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	Immediate	10.79	10.32	.47	.68	.2	2.5 : 1
II	Immediate	7.96	7.45	.51	.34	.5	11 : 1
III	Immediate	8.78	8.24	.54	.23	.8	75 : 1
IV	Review	12.89	5.06	7.83	.62	4.1	.....
V	Review	15.33	12.62	2.71	.69	1.4	20,000 : 1
VI	Review	19.33	16.93	2.40	1.30	.7	38 : 1
VII	Review	41.13	39.94	1.19	1.78	.2	2.5 : 1

TABLE II. RESULTS OF EXPERIMENTS BASED ON SCORES ON TRUE-FALSE TESTS

Experiment	Kind of Value Measured	Net Score with Notes[a]	Net Score without Notes	Difference in Favor of Notes	Standard Deviation of the Difference	Experimental Coefficient	Approximate Chances That the Real Difference Is above Zero
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
II	Immediate	16.90	17.07	−.17	.27	−.2	2.5 to 1
III	Immediate	22.52	22.85	−.33	.42	−.3	3.9 to 1
V	Review	41.39	39.95	1.44	.57	.9	160 to 1
VI	Review	37.00	38.00	−1.00	2.05	−.2	2.5 to 1

[a] *Net score* here means the sum of the means as per McCall, *op. cit.*, p. 187.

[b] Probably several million to one, but not given in McCall's table.

experimental coefficient is dependent upon the number of subjects, those experiments having only a few subjects have low coefficients even though the numerical differences in scores may be greater. A careful analysis of all the facts seems to justify the conclusion that the immediate value of notes is less than the delayed-review value. This immediate value is of sufficient importance, however, to justify the practice of taking notes, even if there be no opportunity to use them later.

The extremely low coefficient, 0.2, for Experiment VII is probably due to one, or possibly both, of two causes: (1) The class periods were devoted to recitations on the textbook rather than to the presentation of new subject-matter. (2) The examination period of one hour was too short to permit the students to finish their papers, and the test may thus have failed to measure differences which actually existed. The writer prefers the former view.

The results as measured by the true-false test afford an extremely interesting contrast to those just examined. It will be recalled that in four of the experiments true-false tests and the general quiz were given during the same period; they constitute, therefore, supplementary or additional measures by a different type of testing device. The essential difference between the two types of the tests seems to be that the true-false test measures ability to recognize specific points, whereas the general quiz measures the ability to recall a number of points and organize them into their natural sequence. Table II shows the results of the four experiments in which the true-false tests were given.

Curiously enough according to Table II three of the experiments yielded negative coefficients. The only positive coefficient, namely, that of 0.9 obtained by Experiment V, is lower by .5 than the coefficient obtained from the general quiz in the same class. The fact that none of the negative coefficients is large enough to be significant is worthy of note and admits of the possible interpretation that the true-false tests are really not satisfactory measures of anything and that the negative coefficients obtained are merely due to chance. The coefficient of 0.9 in

Experiment V contradicts this interpretation, however, and the fact that there are no slight positive coefficients also makes it somewhat doubtful.

The writer's interpretation of the contrast between the results obtained from the general quiz and the true-false test is that the latter measures a different quality of student achievement. It measures ability to *recognize* specific points, whereas the general quiz measures ability to *recall and organize* a series of points. Taking notes seems to be the more effective method of acquiring the latter ability, and listening without taking notes seems to be slightly more effective in acquiring the former. Indeed, this is what we might expect, since the note-taker, because of his greater attention to the important points, may be somewhat inclined to overlook and entirely miss certain statements of minor importance which may later be included in the true-false test.

The fact, that the negative coefficients are low and the one positive coefficient high, suggests that in so far as these results may be of practical guidance to students they should encourage rather than discourage note-taking. This is due to two facts: (1) Notes are taken for their review value as well as for their immediate value, and even the true-false test shows undoubted review value in Experiment V. (2) Efficient study demands a much greater degree of ability to select, organize, and recall important facts than of ability to recognize specific points when they are stated.

Summary and Conclusions

We have now examined the results of seven experiments, involving a number of classroom situations typical of those in which students ordinarily take notes. These have included a sufficiently large sampling of subjects to be fairly representative of college students in general. On the basis of these results we are justified in concluding:

1. When results are measured by a general quiz of the traditional type immediately after the lecture the note-takers show

a fairly high degree of superiority over those who do not take notes.

2. When results are measured by the general quiz after a period of days or weeks and after there has been opportunity to review the notes which were previously taken, the note-takers show marked superiority over those who do not take notes.

3. When results are measured by the true-false test immediately after the lecture, the note-takers are slightly inferior to the listeners.

4. When results are measured by the true-false test after a period of days or weeks, involving the opportunity to review the notes, the note-takers are at an advantage, but not so great as in the case of the general quiz.

5. The difference in results between note-taking and listening is not only characterized by the amount of the lecture retained, but probably also by the quality of the organization of the ideas retained.

6. The value of taking notes depends upon the nature of the class period and is greater when the hour is given to the presentation of new subject-matter than when devoted to recitation on the textbook.

Additional investigation to supplement the results reported in this article might well include studies of the following problems:

1. Further analysis of the qualitative differences in the results of note-taking and of listening.

2. Further investigation of the true-false test as compared with the general quiz as a measuring instrument.

3. Determination of the results of different methods of taking and using notes.

4. Determination of the results of note-taking under other circumstances, for example, by high-school students.

5. Determination of the results after periods of training in the use of notes or in depending on memory instead of notes.

6. Determination of the value of taking notes or making outlines of material in books.

Editorials

THE DIFFICULTY OF THE NUMBER COMBINATIONS

The appearance in the *Journal of Educational Psychology* for 1925 of another article on the difficulty of number combinations (Batson and Combellick) leads us to raise the question, What use can we make of the results of such studies?

Holloway began it in 1914, and in some respects no one has done a better job than he. To be sure he confined his attention merely to addition and multiplication—a limitation which, while we may regret it, is nevertheless defensible. If writers of doctors' dissertations would more rigidly limit their problems! But that's another story. Another way in which Holloway circumscribed his work—and this time one can't be so forgiving—was in reporting direct and reverse facts as the same facts. For example, in his study, $6 + 2$ is the same as $2 + 6$; 7×4, the same as 4×7. Holloway's own figures must have shown him that this was not the case, but in other respects his study is admirable. It not only breaks new ground, but it does so impressively. It employs two measures of difficulty, namely, learning and retention after learning, that have received little subsequent attention from investigators. Moreover, it definitely starts the analysis of textbooks when it shows the number of times the fundamental addition facts are repeated in three "number primers."

Counts in 1917 gave some attention to the difficulty of number combinations although this was not his main problem. His data were gathered from the administering of the Cleveland Survey Test. Limited as he was to the offerings of this particular test, he was able to give only a fragmentary report on the number facts.

Smith in 1920 introduced "time of response" as a measure of difficulty although, except in addition, he also reported the number of errors. His tables cover all four operations and dis-

tinguish—again with the exception of addition—*all* the basic facts instead of only forty-five of them.

Clapp in 1924 presented a complete report in which he not only recognized all the basic facts in each operation but also reported his facts separately for each grade from the third to the eighth inclusive. It is worth noting that he gave percents of error not only when the isolated combinations were presented but also when the same combinations were required in the more advanced processes.

Now comes the article by Batson and Combellick. It gives for each of the one hundred facts of addition and multiplication the reaction time of adults.

Thus we have from several points of view data on the difficulty of the number facts. In their broad outlines these data tell much the same story. Clapp and Holloway on the forty-five addition facts common to their studies correlate 0.67 in their determination of difficulty in spite of considerable difference in method. On multiplication facts they correlate 0.84. Smith and Clapp correlate 0.78 on the difficulty of addition facts, although the former used time reaction and the latter the percent of errors as measures of difficulty. The same investigators show correlations for division of 0.40, for subtraction of 0.71, and for multiplication of 0.79. Other correlations, although they are literally too numerous to mention, are likewise strong.

The question persists, What is the best use we can make of these facts? Thus far the suggestion most frequently offered is to present to the children opportunities for practice in accordance with the known difficulty of the number facts. Following this suggestion after it was first made by Holloway, not a few investigators have scored up the occurrences of each combination in arithmetic textbooks. The result has been to the general effect that difficult facts appear much less frequently than easy facts. Accordingly, the effort is being made to provide a better balanced drill, both in textbooks and in practice material.

More than this is possible, however. Those who argue that certain number facts are hard, *because* they have not been suffi-

ciently practiced, are evidently only partly right. The inverse correlation between frequency of occurrence in textbooks and difficulty is far from perfect. On the basis of Clapp's data the coefficient averages about —0.42. If the relative emphasis on practice as afforded by textbooks is a reasonably close measure of the relative emphasis afforded by the school, then there would appear to be different degrees of difficulty *inherent in the combinations themselves*. It appears likely that $7 + 9$ is harder by its very nature than $2 + 2$. If this is the case, something besides the more frequent repetition of $7 + 9 = 16$ is suggested.

This brings us to the initial teaching act. If $7 + 9$ is hard, it is good practice to introduce it as we do other things that are hard. In the first place, we may do well to postpone it until easier facts have been taught—facts which lead up to and support this particular fact. Do we not do this in other subjects? Who teaches hard words first in spelling? Does any one teach division by using two-place divisors before one-place divisors? In reading do we not begin with short sentences and easy words? We feel instinctively that this is right. Then let us in respect to the number combinations, adopt an order of teaching which takes account of the probable inherent difficulty of the several combinations.

In the second place, let us teach the difficult facts with the greater care which their difficulty requires. If we are truly expert teachers we shall know how to do this. We shall have more frequent resort to concrete material in the case of a hard combination. We shall postpone longer the moment when the pupils are left without support to make their own response. We shall dwell longer upon applications to meaningful situations. And we shall take advantage of number facts already presented.

A word—though not more than a word—on this last point, $7 + 8 = 15$ is a hard fact, one which possibly a third of the children of grade III will get wrong. But $7 + 7 = 14$ is an easy combination. Apparently only about one-twelfth of the third-grade children get it wrong. Now, if the combinations have been taught in an order which takes even slight account of their dif-

ficulty, $7+7=14$ will have been presented and extensively practiced before $7+8=15$ is reached. This fact permits the older association to be brought to the aid of the new one. We may show that, since 8 is one more than 7, $7+8$ must be one more than $7+7$.

Finally, knowledge of the relative difficulty of combinations will enable the teacher intelligently to supplement even the improved textbooks and practice materials which modern research is likely to bring forth. There are reasons why it may never be possible to give in a textbook practice on number combinations in strict conformity to their difficulty. It may not even be desirable to do so. Consider, for example, the preponderance of 1's involved in borrowing or in equal additions in the process of subtraction. This is only one of many considerations which may make it inadvisable for a good textbook to parallel the difficulty ranking of number combinations.

Much, therefore, depends upon the teacher; and in order to do her part, she must know rather accurately the relative difficulty of the combinations. If she has this knowledge there are many possibilities. For example—to apply in a very general way the conclusions of the investigators—she may, in conducting classroom games and devices, avoid both 1 and 2 as addends or factors; make only sparing use of 3, 4, and 5 in combinations with numbers not over 5; introduce rather frequently 3, 4, and 5 in combination with addends or factors larger than 5; and bring in most frequently the pairs of addends or factors in which 6, 7, 8, or 9 is combined with any one of these four numbers. There are obviously 16 of these last pairs. Considered as addends, they yield 16 addition facts and 16 corresponding subtraction facts; considered as factors they yield 16 multiplication and 16 division facts (to say nothing of 16 partition facts of the type 1/6 of 42).

But a teacher may use knowledge of the difficulty of number facts in still other ways. Repetition is not the sole law of learning. We have already spoken of the desirability of richer and more numerous associations during the early presentation of

the more difficult number facts. Somewhat the same idea may be carried over into the drill. Each occurrence of a hard combination should be singled out for attention. It should be something more than an incident; it should be an event. Thus, each difficult number fact should not only occur frequently, it should occur vividly and impressively.

Thus there will always be need for a varied classroom procedure. Here is the teacher's opportunity to supplement the textbook and to do so in conformity to the difficulty of the elements involved. In order to accomplish this she must know something about the difficulty of these elements. Sometimes we become impatient at the slowness with which discoveries fundamental to the expertness of the teacher percolate into the curriculum of the training school. For our part we should far rather have a prospective elementary-school teacher know the placement of number facts in four or five difficulty groups than to have her grounded in the ratio idea or the Grube method, or even to be able to do pretty things with splints, shoe pegs, and colored yarn.

B. R. B.

A PROPOSAL FOR A BILL

The text of the new education bill as it is to be presented to Congress on the opening day of the session is here included. The draft has been prepared by the Legislative Committee of the National Education Association and the representatives of other national organizations which favor a national Department of Education.

To Create a Department of Education and for Other Purposes

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled, That there is established at the seat of Government an executive department to be known as the Department of Education, which shall be under the control and direction of a Secretary of Education to be appointed by the President, by and with the advice and consent of the Senate. Section 158 of the Revised Statutes is amended to include the Department of Education and the provisions of Title IV of the Revised Statutes, including all amendments thereto, shall be applicable to the department. The Secretary of Education shall cause a seal of office to be made for the Depart-

ment of Education of such device as the President shall approve, and judicial notice shall be taken thereof.

SECTION 2. There shall be in the Department of Education an Assistant Secretary of Education, to be appointed by the President, and to receive a salary equal to that of any other assistant in the Department. The Assistant Secretary shall perform such duties as may be prescribed by the Secretary of Education or required by law. There shall also be a solicitor, a chief clerk, and a disbursing clerk, and such chiefs of bureaus and such scientific, technical, and clerical assistants as may be necessary to carry out the provisions of this act and as may be provided for by Congress from time to time.

SECTION 3. (a) The Bureau of Education and all pertaining thereto is transferred from the Department of the Interior to the Department of Education.

(b) The office of Commissioner of Education is abolished, and the authority, powers, and duties heretofore conferred and imposed by law upon the Commissioner of Education shall be exercised and performed by the Secretary of Education.

(c) The Federal Board for Vocational Education is transferred to the Department of Education, and all the authority, powers, and duties heretofore conferred or imposed by law upon the Federal Board for Vocational Education shall be exercised and performed by the Board as a division of the Department of Education. The Secretary of Education shall be a member of the Federal Board for Vocational Education and ex-officio chairman of said board.

(d) The authority, powers, and duties conferred and imposed by law upon the Secretary of the Interior with relation to the Columbia Institution for the Deaf and Howard University, shall be exercised and performed by the Secretary of Education. The Assistant Secretary of Education shall be ex-officio a director of the Columbia Institution for the Deaf, in addition to the directors whose appointment has heretofore been provided by law.

SECTION 4. (a) Except as otherwise provided by this act, all authority, powers, and duties held, exercised, and performed by the head of any executive department in and over any bureau, office, or branch of the Government which is by this act transferred to the Department of Education, or which is abolished by this act and its authority, powers, and duties transferred to the Department of Education, or in and over any business arising therefrom or pertaining thereto, or in relation to the duties performed by and authority conferred by law upon such bureau, office, or branch of the Government, whether of an appellate or revisory character or otherwise, shall be vested in and exercised and performed by the Secretary of Education.

(b) All orders, rules, regulations, and permits or other privileges, which have been issued or granted by any bureau, office, or branch of the Government which is transferred under the provisions of this act to the Department of Education, or which is abolished and its authority,

powers, and duties transferred to the Department of Education, and which are in effect at the time of the transfer, shall, after the transfer, so far as they are not in conflict with the provisions of this act, continue in effect to the same extent as if the transfer had not occurred.

SECTION 5. All officers and employees employed in or by any office, bureau, or branch of the Government, transferred in accordance with the provisions of this act to the Department of Education, are transferred to the Department of Education without change in classification or compensation; and the records and property (including office equipment) are transferred to the Department of Education.

SECTION 6. The Secretary of Education shall have charge in the buildings and premises occupied by or assigned to the Department of Education, of the library, furniture, fixtures, records, and other property pertaining to the Department or hereafter acquired for use in its business. Until other quarters are provided, the Department of Education may occupy the buildings and premises occupied by the bureaus, offices, and branches of the Government which are by this act transferred to or included in the Department of Education.

SECTION 7. In order to coordinate the educational activities carried on by the several executive departments, and to devise ways and means of improving the educational work of the Federal Government, there is hereby created The Federal Conference on Education which shall consist of one representative and one alternate appointed by the head of each department. The Federal Conference on Education shall not report as a body to any one department, but each representative shall report the findings of The Federal Conference on Education to his own department for consideration and independent action.

SECTION 8. (a) The Department of Education shall collect such statistics and facts as shall show the condition and progress of education in the several States and in foreign countries. In order to aid the people of the several States in establishing and maintaining more efficient schools and school systems, in devising better methods of organization, administration, and financing of education, in developing better types of school buildings and in providing for their use, in improving methods of teaching, and in developing more adequate curricula and courses of study, research shall be undertaken in (1) rural education; (2) elementary education; (3) secondary education; (4) higher education; (5) professional education; (6) physical education, including health education and recreation; (7) special education for the mentally and educationally handicapped; (8) the training of teachers; (9) immigrant education; (10) adult education; and (11) such other fields as in the judgment of the Secretary of Education may require attention and study.

(b) The Department shall make available to educational officers in the several States and to all other persons interested in education, the results of the research and investigations conducted by it, and the funds appropriated for printing and binding for the Department of Education

shall be available for the printing and binding of the results of such research and investigations.

SECTION 9. For the fiscal year ending June 30, 1926, and annually thereafter, the sum of $1,500,000, or so much thereof as may be necessary, is hereby authorized to be appropriated, out of any money in the treasury not otherwise appropriated, to the Department of Education, for the purpose of paying salaries and the conducting of studies and investigations, the paying of incidental and traveling expenses incurred in connection with the investigations or inquiries undertaken by the Department and for law books, books of reference and periodicals, and the paying of rent where necessary, and for such other purposes as may be necessary to enable the Department of Education to carry out the provisions of this act. All unexpended appropriations which shall be available at the time when this act takes effect in relation to the various bureaus, offices, and branches of the Government which are by this act transferred to or included in the Department of Education, or which are abolished by this act, and their authority, powers, and duties transferred to the Department of Education, shall become available for expenditure by the Department of Education and shall be treated the same as if such bureaus, offices, and branches of the Government had been directly named in the laws making the appropriations as part of the Department of Education.

SECTION 10. The Secretary of Education shall annually, at the close of each fiscal year, make a report in writing to Congress giving an account of all moneys received and disbursed by the Department of Education and describing the work done by the Department. He shall also from time to time make such special investigations and reports as may be required of him by the President or by either House of Congress or as he himself may deem necessary and urgent.

SECTION 11. This act shall take effect thirty days after its passage, except that the provisions of this act in relation to the transfer of any agency from the jurisdiction and control of one officer to the jurisdiction and control of another, or in relation to the transfer of authority, powers, and duties from one officer or agency to another, shall take effect July 1, 1926.

Reviews and Abstracts

FRANKLIN, EDWARD E. *The Permanence of the Vocational Interests of Junior High-School Pupils.* Baltimore, Maryland, Johns Hopkins Press, 1924. 59 pp. (Johns Hopkins Studies in Education Monograph, No. 8)

The purpose of Dr. Franklin's study is best stated in his own words: "To ascertain how permanent and how reliable are the vocational interests of pupils entering the junior high schools. The enduring interests of the child are the key to his success in life. But when does he acquire, or develop, these interests? If he already has them when he enters the junior high school, the basis for the work of that institution is presented. If he does not have them, it is the duty of the school to aid him in developing them."

The entering pupils (grade VII-B) in eight of the ten junior high schools of Baltimore, are included in the study. In December, 1922, the pupils filled out a questionnaire "devised for the purpose of studying the vocational, recreational, subject-matter, and work interests of the pupils." Three follow-up questionnaires were given: the first in May, 1923, a second one in October, 1923, and a third in December, 1923, one year after the original study. The follow-up cards were then compared with the original ones for the purpose of determining the permanence of choices. Dr. Franklin says in his conclusion: "The vocational interests of junior high-school pupils show a high degree of permanence over a period of at least one year—and that a critical one. Two children out of three have the same preference at the end of the period as they had at the beginning."

The writer feels that before conclusions can be drawn concerning the value of pupils' vocational choices, studies will have to be made upon a much larger scale and over a decidedly longer period of time. Methods will have to be developed which will determine with some certainty the vocational interests which are behind the various vocational choices which pupils make, and which will ascertain more accurately the relationship between vocational interest and ability.

H. B. ALBERTY

Ohio State University

BERRY, JAMES B. *Teaching Agriculture.* Yonkers-on-Hudson, New York, World Book Company, 1924. 230 pp.

One who infers from the comprehensive title of this book that it is designed to cover the entire range of instruction in those fields of knowledge related to agriculture is in error. The author has wisely limited himself instead of attempting to treat the teaching of agriculture in the elementary school, the high school, the vocational school, the college, and the graduate school. He has endeavored to bring together from educational theory, from his own ex-

perience, and from the experience of others, those materials that will be helpful in the teaching of vocational agriculture on the secondary-school level.

Mr. Berry has had considerable experience in teaching, and has had an opportunity, as the county supervisor of agriculture in Crawford County, Pennsylvania, to follow closely the teaching of others. He has a background which makes it possible for him to provide many concrete illustrations.

In his emphasis on seasonal sequence, the importance of the home project, the recognition of community conditions in the determination of what is to be taught, and the use of the farm survey, Mr. Berry will have the support of most of the students of the problems of teaching agriculture on a vocational basis in the secondary school.

The author fails to harmonize his suggested teaching procedures with what he formulates as his fundamental conception of the aim of teaching, namely, "to develop the ability to think purposefully." He says: "In his supervised study of an enterprise, the pupil follows a project-study outline furnished by the teacher." For years many successful teachers have recognized the large opportunity "to think purposefully" that is offered to pupils in developing their own project outlines.

He states that: "The purpose of the home project is to demonstrate the pupil's ability to solve successfully the problems of scientific agriculture arising in connection with an enterprise of the home farm, under life conditions with which he is familiar." In the view of many the primary purpose of the home project is not "to demonstrate" but rather to call forth such "purposeful thinking" as will give the learner a large measure of control of a significant body of skills, knowledges, and attitudes related to farming.

The teacher who is just beginning his service will find many specific suggestions of a helpful character. He is also very certain to be encouraged in his efforts by the author's enthusiasm.

GEORGE A. WORKS

Cornell University

HARAP, HENRY. *The Education of the Consumer—A Study in Curriculum Material.* New York, Macmillan Company, 1924. 360 pp.

When the functional approach to the curriculum is used, an analysis of activities becomes immediately necessary in order to determine the topics of instruction. This analysis may be positive and deal with duties, vocabularies, routine actions, and so on, or it may be negative and list the difficulties, problems, errors, and the like. Some material lends itself to one form of analysis, while some more easily yields to the other. The negative analysis—a listing of problems, errors, or difficulties—is always a shortcut to the positive.

Mr. Harap has adopted in principle a third and different technique. He makes an indirect survey of the needs of consumers of clothing, food, fuel, and shelter. That is, he has selected from multitudinous sources quantitative data concerning the consumption of these materials. These he compares with approved standards of living, and from the comparison of consuming habits and consuming standards he deduces the problems which should concern the curriculum builder.

The characteristic of the work which first impresses the reviewer is the copiousness of the sources studied in the search for quantitative data bearing upon consumption. The following represent only a small section of the literature studied: The Census Reports, the U. S. Bureau of Agriculture yearbooks, Raymond Pearl's summaries of food consumption for seven years, U. S. Housing Commission reports, the monthly data of U. S. Bureau of Labor Statistics, the bulletins of the U. S. Tariff Commission, many reports of private trade agencies, trade journals, and the publications of many semi-public agencies like the Russell Sage Foundation and the National Bureau for Economic Research.

Mr. Harap has performed a valuable service for all those interested in educating the consumer. The mere collection of all this material is in itself a contribution to the raw material of the curriculum. It is of obvious value to the home-economics teacher. Its presence, however, points clearly to the value of such material for a "consumer's course" to be required of both boys and girls in the high school. We firmly believe that training in production—vocational education—is necessary. Why is not training for consumption equally necessary? The difference between culture and crudeness is not a matter of how much we purchase but of how wisely we purchase.

Mr. Harap has made wide and constant use of quantitative material, and in this lies the chief excellence of his text, but a careful examination of the material reveals two points at which his methods are subjective.

In the first place, from many thousands of pages of reports he has selected perhaps one hundred and fifty pages of material, and he does this without indicating in any specific manner just what his basis of selection is. He merely says "the method here is to gather all the evidence which shows what the people actually do in the process of consuming food, clothing, shelter, and fuel." (p. 5.) But obviously this evidence cannot be compressed into some one hundred and fifty pages. One is left with the feeling that the author must have been unconsciously guided by his conception of certain problems and that some other investigator might have selected different material. This feeling of uneasiness might have been removed if some objective basis of selection had been used and described.

In the second place, the lack of objective criteria seems to the writer also to be present to a slight degree in the deduction of the objectives from the habits and standards of consumers. The objectives are often wider in scope than the material warrants if the author were strictly to follow the statement made on page 10 "the quantitative comparison between modes of living and standards of living yields conclusions which become objectives of educational activity." One feels that here again the author had unformulated and unconscious notions of what should be the objectives of each section (in part antedating the study and in part suggested by it) and added them to the objectively derived objectives. That he recognizes this is of course clear from his statement on page 9 that "usually" the comparisons were quantitatively determined. It would have been better, however, to have indicated, if possible, the supplemental rules which governed their addition.

Mr. Harap has labored diligently to produce a fine piece of work for the school curriculum. He has made, not the fully respectable two blades of grass grow where only one grew before, but a whole field of grass to grow on a barren plot in the educational acres.

W. W. CHARTERS

University of Pittsburgh

SHARP, RUSSELL A. *Teaching English in High Schools.* Boston, Houghton Mifflin Company, 1924. 155 pp. (Riverside Educational Monographs)

A complete tour of Europe in two months! One day spent in Rome! Thirty-six beds in twenty-four nights! Such is the suggestion arising to the reader as he considers this recent addition to the well-known series of Riverside Educational Monographs. One hundred and fifty-five small pages not only suffice for a discussion of Teaching English in High Schools but allow time and space for a preliminary survey of the requisites of a teacher and for a final excursion to hear a brief lecture on modern dancing. One cannot help contrasting this little book with its brother volume, Professor Leonard's *English Composition as a Social Problem* which deals very specifically and almost definitely with one carefully selected, important problem in the teaching of the vernacular.

Despite this limitation, however, despite the fact that one looks in vain for a discussion of the very vital and insistent question of improving the pupils' everyday speech, *Teaching English in High Schools* is an excellent book. Nothing very revolutionary; nothing very startling; for the writer evidently realizes that progress here demands no great upheaval, no educational cataclysm, but that rather will it come largely from doing more effectively those things we are already attempting, by pressing on gradually but steadily from point to point and consolidating and fortifying the ground gained.

The author is closely in touch with the advances which have marked the teaching of English in the past decade; he has discussed the most important of these with a gratifying sanity and clarity. Indeed one feels on every page, the impress of the wide-awake, progressive, thoughtful teacher drawing upon the stores of an extended and valuable experience in the classroom. Perhaps the most commendable feature of the book is to be found in the many excellent suggestions for classroom procedure which freight its pages. Another almost equally excellent feature is the brief summaries of the advantages and the disadvantages of such matters as the project, the socialized recitation, and other innovations which have, of late years, been gaining a foothold in the English class.

The author has conducted us on a helpful, stimulating, suggestive trip across the field of English. "Home-keeping youths have ever homely wits"; so teachers of English and high-school administrators should by all means profit by the many valuable hints and the sound judgment shown in this little book.

H. G. PAUL

University of Illinois

BOWMAN, CLYDE A. *Graphic Aids in Occupational Analysis for Guidance and Teaching.* Bruce Publishing Company, Milwaukee, Wisconsin, 1924. 100 pp.

This book has many of the virtues and a few of the disadvantages of a pioneer book. Certainly those interested in progressive methods in vocational education and guidance will give it careful attention and many workers will try out its methods in the classroom.

We have had occupational specifications of many sorts, but none to the knowledge of the reviewer which have tried to put in graphic form the complete story of the divisions and subdivisions of a trade or a job, at least in such elaborate form as has been worked out by Mr. Bowman and his colleagues. Organization charts of industrial and commercial enterprises are quite common: the present book applies the same method to auto mechanics, sheet metal, forging, and dozens of other kinds of work. The textual material accompanying the charts is an aid in understanding them, but in the opinion of the writer there is not enough help to the teacher in telling just how the charts are to be used with the children.

It is very questionable whether classes in occupations need such detailed analysis as are herein elaborated. What is rather needed is a clear diagram of the several kinds of workers. There is possibly in this pioneer book an attempt to spread over too much territory.

The psychologist will derive help from a study of this book since it is one of the first attempts to use the graphic method extensively. It is possible that the book is more valuable for children who have learned to visualize, than for pupils who prefer other methods of study.

JOHN M. BREWER

Harvard University

BAILOR, EDWIN M. *Content and Form in Tests of Intelligence.* New York, Teachers College, Columbia University, 1924. 74 pp. (Teachers College, Columbia University, Contributions to Education, No. 162.)

This is a contribution to theoretical educational psychology in the field of intelligence measurement, particularly as to the validity of the "general factor" theory of intelligence. The author defines *content* from the point of view of this study, as consisting specifically of three kinds: (1) verbal, (2) numerical, and (3) spatial. He defines "form" as the objective mold in which the test is cast. His main argument is based on the inter-correlations of tests of 1,039 white boys of high-school age, given 15 tests in 1922, and repeated one year later.

The large number of cases, the careful selection of subjects as to age, sex, color, residence, and repetition of all tests one year apart, make possible the use of such refined statistical analyses as corrections for attenuation, and the

use of partial coefficients computed from the corrected coefficients. Correction for attenuation rules out statistically the factor of different length—of the time limit of the test. In the larger group the resulting difference in form is shown by three coefficients: the corrected coefficient of correlation between numbers and words is, .7614; between numbers and space, .7565; and between words and space, .6797. There is thus no evidence to bear out Spearman's contention of a single, general intellectual factor since validation of this theory requires each of the above correlations to be 1.00. All indications suggest basic discreteness among intellectual factors. General ability is thus made up of a number of discrete or special capacities. The pedagogical implication is that there are innate differences in learning rates and limits of acquirement of differing educational contents.

The practical conclusion of the research for test construction is that in constructing a scale for measuring any general ability, each test must be weighted, not equally or arbitrarily, but according to an empirical law, weighted with a differential regression weight. The need for special-abilities tests, even in the measurement of intelligence, is evident.

It is hoped that the author may see fit to publish at an early date all the intercorrelations, means, and standard deviations of the two givings of the 15 tests. These, because of the low probable errors and repetition of all tests, would be of inestimable worth for future theoretical studies of many types. Contrary to one's expectations, from the titles no light is shed upon the problem of the form of test—analogies, completion, multiple choice, true-false—in which a given content, for example, arithmetical reasoning problems, should preferably be cast. The author points out the need for further research in this field.

HERBERT A. TOOPS

Ohio State University

PACKER, PAUL C. *Housing of High-School Programs.* New York, Teachers College, Columbia University, 1924. 51 pp. (Teachers College, Columbia University, Contributions to Education, No. 159)

"Scientific data which will help superintendents and architects plan satisfactory high-school buildings are very much worth while. A modern school building should be planned to fit courses of study which will be offered in a given community. School work should not be forced to conform to the physical plant, as is too often the case."

Dr. Packer's study is primarily devoted to the development of a technique which will determine the number of instructional rooms required for the program of studies in any comprehensive high school. A brief sketch is given showing how high schools have grown in numbers, size, and curricula. Several examples are given to show how programs are now being housed in high-school buildings. The study indicates that there is need for more careful planning.

An analysis of high-school programs was also made to determine the number of rooms which should be included in a high-school building. Four

elements are given as necessary in order to translate a given program of studies into the number of rooms required: (1) the registered number of students in the subject, (2) the average number of daily periods devoted to the subject, (3) average size of class, and (4) the number of periods in the day. These four factors are used as a means of determining the needed number of rooms in the following formula:

$$\frac{\text{Factor 1} \times \text{Factor 2}}{\text{Factor 3} \times \text{Factor 4}} = \text{Number of rooms}$$

Several applications of the formula are made, but more detailed explanation would be of benefit to the average school superintendent.

P. R. STEVENSON

Ohio State University

WILLET, GEORGE W. *The Public-School Debt in Illinois.* New York, Macmillan Company, 1924. 97 pp.

Speaking generally, this study describes and interprets the trends of bonded indebtedness for schools in Illinois from 1873 to 1922. Speaking more specifically, it discusses the location of bonded indebtedness, the causes for increased bonded indebtedness, the outlook for future bonding, the floating of bond issues, the cost of bonded indebtedness, floating indebtedness, and school indebtedness compared with that of other taxing units, and the probability of repayment of the indebtedness.

Some outstanding findings of the study are: bonded and floating indebtedness have increased very rapidly, particularly in recent years; indebtedness has increased much more rapidly than the total population and school attendance; school indebtedness has increased faster than other indebtedness; some districts have many times as much indebtedness as others; the high schools, and particularly the smaller ones, have far more per pupil indebtedness than the elementary schools.

Some of the outstanding merits of the study are: it is timely; it is unusually clearly and interestingly written, particularly for an academic document; and it bears the chief earmarks of a scientific document in that it is accurate, objective, impartial, and verifiable.

Some things missing that, if present, would add to the value of the study are: more interpretations of facts; more data on trends of indebtedness compared with trends of real or assessed wealth; some data on the value of school property compared with school indebtedness, that is, assets compared with liabilities; and a summary.

The reviewer believes the study the best treatise that has yet appeared on the public-school debt of a state. It will be of interest primarily to the people of Illinois, but because of its generally applicable technique it will also be of interest to students of school finance everywhere.

WARD G. REEDER

Ohio State University

RUCH, G. M. *The Improvement of the Written Examination.* Chicago, Scott, Foresman and Company, 1924. 193 pp.

Although there has been much recent interest in the improvement of examinations, the reviewer believes that this is the first volume entirely devoted to the topic. The author states "that the function of this book is to suggest a methodology of examination construction which is immediately serviceable to the classroom teacher or supervisor." He deals with the functions of written examinations, the criteria for judging them, the sources of error, and the types and construction of the newer objective examinations. He also discusses a few experimental data and devotes a chapter to pertinent statistical procedures.

The experiment in which the author compared several types of new examinations is interesting and affords valuable data. One regrets, however, that it was not more extensive and also that it did not deal with the completion and the matching types of tests. Ruch concludes that the recall and the multiple-response types "seem to be superior in every point of view" to the true-false. The reviewer believes, however, that since their superiority over the true-false test was not great it is probably wise to use the latter occasionally for the sake of variety.

The author states that the newer types of examinations may be so constructed as to require reasoning as well as mere factual memory, yet suggests that the speed of high-school or college students on a recall test should be four to eight items per minute; on a recognition test, six to ten; and on a true-false test, ten to fifteen. Certainly at these speeds little reasoning can be done.

Several statements of doubtful accuracy are to be found. For example, Ruch states that "measurement is always sampling," again "that an examination of ten questions, or worse still, five questions . . . . is a very small sampling is evident." Although both statements express general truths the reviewer does not believe that they always hold. In a limited field one can secure a complete measure of ability. For example, a test may include all the addition combinations of single digits. Likewise ten or even five questions may give more than a very small sampling if they are topical or in some other way call for large amounts of material. In the chapter on statistical considerations several statements appear which might mislead the tyro. For example, it is stated without any qualifications that "the standard deviation includes 34 per cent of the total distribution." This is only true for a normal distribution and also in case the standard deviation is laid off from the mean.

Despite its omissions, inaccuracies, and other faults, the reviewer predicts that this volume will be widely read and will do much to stimulate the movement for better examinations.

C. W. ODELL

University of Illinois

SUZZALO, HENRY. *Our Faith in Education.* Philadelphia, J. B. Lippincott Company, 1924. 108 pp.

As is suggested by the title, this little book is concerned, in the main, with the large social significance of education. A vision of this social sig-

nificance, as the author insists, is necessary to sustain the teacher's faith in his calling. Other things are subordinate to it in importance:

There is a sense in which values are more important than technique. The most skilful mind scarcely starts favorably upon the road of important accomplishments unless it knows the high significance of its errand. It is by seeing the larger relations and dependencies between education and the nobler aspirations of human life that teachers get a footing for their faith in the things they do.

This "high significance" stands out prominently when we observe how every road to the realization of our national faith and aspiration leads towards education. In education we find the promise of progressive improvableness upon which our faith in popular sovereignty is based. Education is indispensable for insuring a "fair chance" and "equal opportunity," and for the training of a national character which will make democracy something more than an empty name. It is only by education that we can develop a society in which leadership and followership will be determined by relation to the well-being of the whole group, and not by the accident of social status or blind circumstance. As measured by the standard of social service our common schools are undergoing significant improvement.

The chapter on "Our Higher Schools" offers both a defense of higher education, on the score of the need for trained leadership, and a statement of the main problems with which the institutions of higher learning are confronted. One of these is to bring within their gates all those who are fit for advanced training. Another is to improve the quality of its teaching, so that "through good teaching every scientist will become a humanist in a fine functional sense" (p. 100). A third problem is to develop a rigid discipline of thinking. "Bad logic in a class discussion should become as intolerable as a discord in a musical studio" (p. 102). Lastly, there is the problem of research. "It is the largest problem before an aspiring university, but its solution waits on the other reforms suggested" (p. 105).

A summary of this sort necessarily falls far short of doing justice to this book. It fails to transmit the author's fine enthusiasm, the suggestion of wider horizons, and the skill of presentation. The book can be cordially recommended to teachers, not only as a professional tonic, but as a means of furthering their understanding of our public schools as "the chief supports of a democratic system of life organized to give liberty and happiness a wider spread."

B. H. Bode

Ohio State University

Publications Received, September 15 to October 15

Aldredge, E. M. and McKee, J. F. *The funny clown and other stories to color; for second and third grades.* Cleveland, Ohio, Harter School Supply Co., 1925. Per set, $.50.

Arnold, Matthew. *Poems.* Selected and ed. by B. L. K. Henderson. New York, Dutton, n.d. 254 pp. $.60.

Barr, A. S. *Elementary school standards for the improvement of teaching.* Ann Arbor, Mich., Edwards Bros., 1924. 159 pp. $1.25.

Barrows, H. H. and Parker, E. P. *Geography; United States and Canada.* New York, Silver, Burdett, 1925. 288 pp. $1.48.

Beeson, C. H. *A primer of medieval Latin; an anthology of prose and poetry.* Chicago, Scott, Foresman, 1925. 389 pp. $2.00.

Bird, J. P. *Essentials of French.* Garden City, N. Y., Doubleday, Page, 1925. 368 pp. $1.50.

Book, W. F. *Learning to typewrite with a discussion of the psychology and pedagogy of skill.* New York, Gregg, 1925. 463 pp. $2.80.

Book, W. F. *The psychology of skill with special reference to its acquisition in typewriting.* New York, Gregg, 1925. 257 pp. $2.00.

Breslich, E. R. *Junior mathematics.* New York, Macmillan, 1925. Book II. 264 pp. $.80.

Brownell, Herbert, and Wade, F. B. *The teaching of science and the science teacher.* New York, Century, 1925. 322 pp. $2.00.

Cather, K. D. *Educating by story-telling.* Yonkers-on-Hudson, N. Y., World Bk. Co., 1924. 396 pp. $2.20.

Clark, Imogen. *Suppose we play; a collection of indoor games for young and old.* New York, Crowell, 1925. 299 pp. $2.00.

Clark, V. B. *Europe: a geographical reader.* New York, Silver, Burdett, 1925. 555 pp. $1.28.

Coffman, G. R. ed. *A book of modern plays.* Ed. for school use. Chicago, Scott, Foresman, 1925. 490 pp. $1.00.

Colum, Padraic. *Three plays: The fiddler's house; The land; Thomas Muskerry.* Rev. ed. New York, Macmillan, 1925. 239 pp. $2.25.

Cooke, E. V. *The garden party: pictures to color, verses to learn and recite.* Cleveland, Ohio, Harter School Supply Co., 1924. 16 pp. $.40.

Corwin, E. S. *The constitution and what it means today.* Princeton, N. J., Princeton Univ. Press, 1924. 130 pp. $2.00

Crabbs, L. M. *Measuring efficiency in supervision and teaching.* New York, Teachers College, Columbia University, 1925. 98 pp. (Contributions to education, No. 175). Cloth, $1.50. Paper, $1.25.

Crowell's handbook for readers and writers: a dictionary of famous characters and plots in legend, fiction, drama, opera, and poetry. Ed. by Henrietta Gerwig. New York, Crowell, 1925. 728 pp. $3.50.

Cubberley, E. P. *An introduction to the study of education and to teaching.* Chicago, Houghton Mifflin, 1925. 476 pp. $2.00.

Dearborn, N. H. *The Oswego movement in American education.* New York, Teachers College, Columbia University, 1925. 189 pp. (Contributions to education, No. 183). Cloth, $1.50. Paper, $1.25.

DeFoe, Daniel. *Robinson Crusoe.* Philadelphia, Winston, 1925. 301 pp. $.80.

DuPuy, W. A. *Our bird friends and foes.* Philadelphia, Winston, 1925. 319 pp. $.80.

Edson, A. W. and Laing, M. E. *Opportunity.* New York, Sanborn, 1925. 296 pp. $.88.

Edson, A. W. and Laing, M. E. *Working together.* New York, Sanborn, 1925. 259 pp. $.84.

Ellis, M. B. *The visiting teacher in Rochester.* New York, Joint Committee on Methods of Preventing Delinquency, 1925. 205 pp. $.75.

The Elson foundation unit; pre-primer work in reading. Chicago, Scott, Foresman, n.d. Manual. $.25. Picture cards. Per set, $.35. Word and phrase cards. Per box, $3.00. Action sentence cards. Per box, $3.50. Silent reading seat work cards. Per set, $.20.

Elson, W. H. and Burris, M. H. *Child-library readers; book eight.* Chicago, Scott, Foresman, 1925. 560 pp. $.96.

Elson, W. H. and Kelly, E. R. *Child-library readers; book three.* Chicago, Scott, Foresman, 1925. 320 pp. $.76.

Elson, W. H. and Runkel, L. E. *Child-library readers; book two.* Chicago, Scott, Foresman, 1925. 240 pp. $.68.

Farr, Alvira. *Puzzle games; drills in addition for second, third, and fourth grades.* Cleveland, Ohio, Harter School Supply Co., 1925. $.40.

Fleeker, O. J. *A school chemistry.* Oxford, Clarendon Press, 1925. 240 pp. $1.25.

Getchell, R. W. *Outline of general chemistry.* Ann Arbor, Mich., Edwards Bros., 1925. 190 pp. $1.50.

Goldstein, F. F. *Pairing words and pictures.* Cleveland, Ohio, Harter School Supply Co., 1924. Per set, $.40.

Graded exercises in Rowe shorthand. Baltimore, Rowe, 1920. 72 pp. $.60.

Hamilton, A. E. *The real boy and the new school.* New York, Boni & Liveright, 1925. 386 pp. $2.50.

Harter's large number cards in strips. Cleveland, Ohio, Harter School Supply Co., n.d. Per set, $.15.

Harter's number cards; for primary number work. Cleveland, Ohio, Harter School Supply Co., 1925. Per box, $.15.

Harter's sight drill cards for number work. Cleveland, Ohio, Harter School Supply Co., 1924. Per box, $.25.

Hazlitt, William. *Essays.* Selected and ed. by C. Calvert. New York, Dutton, n.d. 256 pp. $.60.

Heath, J. F. *Ann's family.* Boston, Lothrop, 1925. 189 pp. $1.50.

Henry, E. G. *The theory and practice of medical social work.* Ann Arbor, Mich., Edwards Bros., 1924. 195 pp. $1.60.

Hessler, J. C. *The first year of science.* Rev. ed. Chicago, Sanborn, 1925. 576 pp. $1.56.

Highsmith, J. A. *Relation of the rate of response to intelligence.* Princeton, N. J., Psychological Review Co., 1925. 33 pp. (Psychological monographs, No. 156). $.50.

Kemmerer, E. W. *The ABC of the federal reserve system.* Princeton, N. J., Princeton Univ. Press, 1924. 215 pp. $1.50.

Laing, M. E. and Edson, A. W. *Busy folk.* New York, Sanborn, 1925. 140 pp. $.72.

Laing, M. E. and Edson, A. W. *Lend a hand.* New York, Sanborn, 1925. 160 pp. $.76.

Laing, M. E. and Edson, A. W. *Neighbors.* New York, Sanborn, 1925. 192 pp. $.80.

Larson, E. L. *One-room and consolidated schools of Connecticut; a comparative study of teachers, costs, and holding power.* New York, Teachers College, Columbia University, 1925. 67 pp. (Contributions to education, No. 182). Cloth, $1.50. Paper, $1.25.

Lentz, T. F. *An experimental method for the discovery and development of tests of character.* New York, Teachers College, Columbia University, 1925. 47 pp. (Contributions to education, No. 180). Cloth, $1.50. Paper, $1.25.

McDonnold, Arline. *Teaching health in public schools.* Temple, Tex., Gresham's, 1925. 84 pp. $1.50.

Manchester, R. E. *The home and school idea in education.* Menasha, Wis., Banta, 1925. 159 pp. $1.00.

Miller, Agnes. *The chimes of Daskam High.* Boston, Lothrop, 1925. 311 pp. $1.50.

Morrison, H. C. *The teaching technique of the secondary school.* Ann Arbor, Mich., Edwards Bros., 1924. 4 vols. $6.85.

Murphy, Marguerite. *Patricia from New York.* Boston, Lothrop, 1925. 316 pp. $1.50.

Navarro y Lamarca, D. C. *Compendio de historia hispanoamericana.* Abbreviated and ed. by R. T. House and Carlos Castillo. Chicago, Scott, Foresman, 1925. 315 pp. $1.60.

Nearing, Scott. *Educational frontiers; a book about Simon Nelson Patten and other teachers.* New York, Seltzer, 1925. 250 pp. $1.50.

New York Society for the Experimental Study of Education. *Contributions to education.* Ed. by J. Carleton Bell. Yonkers-on-Hudson, N. Y., World Bk. Co., 1924. Vol. 1. $2.00.

Owens, C. E. *Principles of plant pathology.* Ann Arbor, Mich., Edwards Bros., 1924. 2 Vols. $4.00.

Persing, K. M. and Myers, G. C. *Summarizing paragraphs; for pp third, fourth, and fifth grades.* Cleveland, Ohio, Harter School Supply Co., 1925. Per set, $.40.

Reigner, C. G. *Applied typing.* Baltimore, Rowe, 1925. 136 pp. $1.40.

Reigner, C. G. *Special instruction and exercises in Rowe shorthand.* Baltimore, Rowe, 1924. 58 pp. $.60.

Rich, F. M. *Projects for all the grades.* Chicago, Flanagan, 1925. 215 pp. $1.25.

Royster, J. F. ed. *American short stories.* Ed. for school use. Chicago, Scott, Foresman, 1925. 342 pp. $.72.

Sample, A. E. *Primary games to teach phonetics.* Chicago, Beckley-Cardy, 1925. 96 pp. $.80.

Sheridan, R. B. *The rivals.* Ed. by John Hampden. New York, Dutton, n.d. 187 pp. $.60.

Simplex daily plan book; arranged for a full year's work. Chicago, Beckley-Cardy, 1925. Unpaged. $.60.

Smith, L. R. *The Pixie-in-school.* Chicago, Flanagan, 1925. 144 pp. $.72.

Smith, A. H. and Cowgill, G. R. *Laboratory directions for physiological chemistry.* Ann Arbor, Mich., Edwards Bros., 1924. 131 pp. $2.00.

Smith, N. B. and Courtis, S. A. *Picture-story reading lessons.* Yonkers-on-Hudson, N. Y., World Bk. Co., 1925. Series I. $.95.

Tickner, F. W. ed. *A book of English odes.* New York, Dutton, n.d. 192 pp. $.60.

Trelawny, E. J. *The adventures of a younger son.* Ed. by E. Collins. New York, Dutton, n.d. 189 pp. $.60.

Tyrrell, J. F. *Tyrrell's American history exercises; true-false exercises, completion exercises, recognition exercises.* Boston, Palmer Co., 1925. Per package, $1.50.

Ward, C. H. *What is English? a book of strategy for English teachers.* Rev. ed. Chicago, Scott, Foresman, 1925. 487 pp. $1.60.

Warshaw, J. and Bonilla, R. H. *The elements of Spanish.* Chicago, Scott, Foresman, 1924. 435 pp. $1.60.

Watson, G. B. *The measurement of fair-mindedness.* New York, Teachers College, Columbia University, 1925. 97 pp. (Contributions to education, No. 176). Cloth, $1.50. Paper, $1.25.

Worst, E. F. *Worst's phonic cards.* Chicago, Flanagan, 1925. Per envelope, 15 cents.

Worst, E. F. *Worst's picture and word cards; for first grade.* Chicago, Flanagan, 1925. Per envelope, $.40.

Worst, E. F. *Worst's primary domino cards; for simple seat work in numbers.* Chicago, Flanagan, 1925. Per envelope, $.20.

Worst, E. F. *Worst's primary number cards; combinations to ten.* Chicago, Flanagan, 1925. Per envelope, $.20.

Zaner and Bloser method writing, based upon arm movement; manual 96. Columbus, Ohio, Zaner-Bloser, 1923. 96 pp. $.15.